S. venezuelae vep rsaA S-lyaer prote

Caulobacter cresce Caulobacter cresce Human Acinus L pro Non-typeable Haemo

Caulobacter S-laye

protein isolated f
H. pylori ORF 04ep
H. pylori outer me
M. tuberculosis im
Mycobacterium tube
M. tuberculosis an
M. tuberculosis an
H. tuberculosis an
H. tuberculosis re
Neisseria gonorrhe
Neisseria meningit
Haemophilus influe
Haman ORFX ORF851

Filamentous haemag Neisseria IgA-Prot Arabidopsis phytoc

Enterococcus faeca Enterococcus faeca Haemophilus paraga Murine T protein f Haemophilus paraga Human cytomegalovi

Enterococcus

mediterranei ri

Perfect score:

Title:

Sequence:

OM protein

on O

Scoring table:

Searched:

Database

```
"given as Xaa in the specification; Lys is deduced from the DNA sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tir; translocated intimin receptor; Hp90; enteropathogenic;
EPEC; infection; diagnosis; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "putative transmembrane domain"
364..386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note= "putative transmembrane domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EPEC E. coli translocated intimin receptor (Tir)
                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'note= "encoded by AAA"
                                                                                                                                                                                                                         .846
                                                                                                                                                                                                                                                                                                                                     AAB10763
AAW56322
                                                                                                                                                                                                                                                407
                                                                                                                                                                                                                                                                                                      AAY00140
                                           AAR48993
                                                                                                                                                                                                                                                                                 AAW79280
                                                                                                                                                                                                                                                                                             AAY00138
                                                                                                                                                                                                                                                                                                                  AAY00142
                                                                                                                                                                                                                                                                                                                            AAW56319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY06220 standard; Protein; 549 AA
                                                                                                                                                                                                    AAB01
                                                                                                                                                                                                                           AAB01
                                                                                                                                                                                                                                                 AAY87
                                                                                                                          AAW8]
                                                                                                                                                                                          AAB0]
                                                                                                                                                                                                                                       AAB4]
                                                                                                             AAY1
                                                                                                                                                                                                                                                                                130
130
130
130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-CA01042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0065130
 (first entry)
2514
1180
1188
1536
1536
2541
3596
3647
741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         234..253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09924576-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-MAY-1999
133.5
132.5
132.5
132.5
131.5
130.5
129
                                                                                                                                                                                                                                               123.5
123.5
123
123
                                                                                       128
128
128
126
126
126
125
124
124
124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY06220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Domain
                                                                                                                                                                                                                                                                                                                                                                                                                                         7
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY06220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Key
  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                               EPEC E. coli trans
EHEC E. coli trans
Intimin C-terminal
EPEC E. coli trans
M. catarrhalis str
Endoglycocaramidas
Haemophilus influe
Haemophilus adhesi
FANCIR3 protein fr
Streptomyces venez
Human ORFX ORF709
                                                                                       (without alignments)
1046.951 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                           ; Search time 31.79 Seconds
                                                                                                                                              549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                      US-09-189-415A-2
2800
1 MPIGNLGNNVNGNHLIPPAP......GETAVSSVNAAPTPGPVRFV
                                                                                                                                                                                                                                                                                                                                  SIDSI/gcgdata/geneseq/geneseqp/AA1980.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1981.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1982.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1982.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1984.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1986.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1986.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1980.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1980.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1980.DAT:*

SIDSI/gcgdata/geneseq/geneseqp/AA1990.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1991.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1992.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1995.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1995.DAT:*

SIDSI/gcgdata/geneseqf/geneseqp/AA1995.DAT:*

SIDSI/gcgdata/geneseqf/AA1995.DAT:*

SIDSI/gcgdata/geneseqf/AA1996.DAT:*

SIDSI/gcgdata/geneseqf/AA1996.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DAT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /SIDS1/gcgdata/geneseg/genesegp/AA2001.DAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDS1/gcgdata/geneseq/geneseqp/AA1997.
/SIDS1/gcgdata/geneseq/geneseqp/AA1998.
ISIDS1/gcgdata/geneseq/geneseqp/AA1999.
/SIDS1/gcgdata/geneseq/geneseqp/AA2000.
           GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                        Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                                    412676 segs, 60623988 residues
                                                                           September 27, 2001, 14:20:24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY06221
AAB20576
AAW68213
AAW68201
AAW10660
AAB23860
AAB23860
AAW196770
AAW19629
AAB40945
                                                      - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY 06220
                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                            A_Geneseq_0601:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2411
2353
3096
4630
1532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match
```

1492 1492 165 149 141 139.5 136.5 136.5

Score

Result No.

(UYBR-) UNIV BRITISH COLUMBIA

```
PQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADP 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                                                                                                   SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALM 420
                                                                                                                                                                                                                                                                                                                                  61 VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                         61 vdsrdipglptnpsrlaaatsetcllggfevlhdkgpldilntqigpsafrvevqadgth 120
                                                                                                                                                                                                                                                                                                                                                               AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTR 180
                                                                                                                                                                                                                                                                                                                                                                       IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL 360
                                                                                                                                                                                                                                                                                                     1 MPIGNICANNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                                                                                                                                                                              New translocated intimin receptor useful for treating infection by enteropathogenic or enterohemorrhagic Escherichia coli
                                                                                                                                                                                                                                                                        99.9%; Score 2798; DB 20; Length 549; 99.8%; Pred. No. 1.1e-210; iive 0; Mismatches 1; Indels 0
             Σ
             Stein
                                                                       Claim 6; Page 55-58; 91pp; English
             Kenny B,
                                                                                                                                                                                                                                                                                          Conservative
              Finlay BB,
                            WPI; 1999-337712/28
                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                     Sequence 549 AA;
                                    N-PSDB; AAX58858
              Devinney R,
                                                                                                                                                                                                                                                                                    Best Local Sim
Matches 548;
                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
g
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

ö

0; Gaps

```
The present sequence represents Tir, a novel translocated intimin receptor (formerly termed Hp90) from an enterohaemorrhagic receptor (formerly termed Hp90) from an enterohaemorrhagic receptor (formerly termed Hp90) from an enterohaemorrhagic secreted and the pathogens are isolated tir polynucleotide (see AX58859). Tir proteins are secreted by attaching and effacing pathogens such as EHEC and EPEC (see AX706220) E. coli. The bacterial pathogens insert their own receptors into mammalian cell surfaces, to which the pathogen then adheres to trigger additional host signaling events and actin adheres to trigger additional host signaling events and actin be performed by use of antibodies that bind to Tir to detect the protein or the use of antibodies that bind to Tir to detect the protein or the use of nucleic acid probes for detection of nucleic acids, Tir peptides, a recombinant method for producing recombinant Tir, and a kit for the detection of antibodies which bind to Tir, and a kit for the detection of antibodica which bind to Tir, and a kit for the detection of antibodica a protective immune response is also provided. With Tir to induce a protective immune response is also provided. In addition, Tir fusion proteins can be used in attenuated E. coli and the provided immune response is also provided. In addition, the binding of bacterial pathogens to their interfere with the binding of bacterial pathogens to their
                 New translocated intimin receptor useful for treating infection by enteropathogenic or enterohemorrhagic Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "encoded by codon of 1 apparent nucleotide, causing frameshift in the DNA sequence"
GYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNAA
                                                                                                                                                                                                                                                                                                                                                   Tir; translocated intimin receptor; Hp90; enterohaemorrhagic; EHEC; infection; dlagnosis; vaccine.
                                                                                                                                                                                                                                                                                                                     EHEC E. coli translocated intimin receptor (Tir).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Stein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 claim 7; Page 55-58; 91pp; English.
                                                                                                                                                                                                          AAY06221 standard; Protein; 559 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Devinney R, Finlay BB, Kenny B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (UYBR-) UNIV BRITISH COLUMBIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-CA01042.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97us-0065130.
                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1999-337712/28.
N-PSDB; AAX58859.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference 453
                                                                               541 PIPGPVRFV 549
                                                                                                                    541 ptpgpvrfv 549
                                                                                                                                                                                                                                                                                                                                                                                                                     Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LO-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09924576-A1
                                                                                                                                                                                                                                                                                      16-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-MAY-1999.
                                                                                                                                                                                                                                                     AAY06221;
         481
                                                                                                                                                                                               AAY06221
                                                                                                                                                                                                                                                   g
                                                                                 δy
                                                                                                                    g
```

Sequence

```
WPI; 1999-337712/28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EPEC E. coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09924576-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UYBR-) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Devinney R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY06213;
                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                               AAY06213
                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
δy
                         12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Intimin; Tir binding domain; inhibitor; intimin adhesion; screening;
Tir-independent eukaryotic cell binding activity; bacterial infection;
                                                                                                                                                                                                                                      61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG 118
                                                                                         119 THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR 178
                                                                                                                           179 TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGI-----DDGV--VSETHTSTTNSSVRS 230
                                                                                                                                                                     177 grilellepkgtg-----eskgageskgvgelresnsgaenttetgtststsslrs 227
                                                                                                                                                                                        DPKFWVSVGAIAAGLAGLAATGIAQALALIPEPDDPTTTDPDQAANAAESATKDQLTQEA 290
                                                                                                                                                                                                   FKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQ 350
                                                                                                                                                                                                                                                                351 HARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTT-----HT 403
                                                                                                                                                                                                                                                                          404 VVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGG---AR 460
                                                                                                                                                                                                                                                                                                                      408 venkpanntpaggnvdtpgsedtmesrrssmastsstffdtss------iggpcrir 458
                                                                                                                                                                                                                                                                                                                                        461 NSLSAHQ-------PEEHIYDEVAADDGXSVIQNFSGSGPVTGRLIG 500
                         Gaps
                                           1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                            1 mpignlghnpnvnnsippapplpsqtdga--ggrgqlinstgplgsralftpvrnsmads 58
                         70;
        Length 559;
                                                                                                                                                                                                                                                                                                                                                                           501 TPGGGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                      Intimin C-terminal Tir binding domain amino acid seguence.
                         Indels
                ; Pred. No. 1.7e-108; 64; Mismatches 133;
       DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dougan G;
       Score 1492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Frankel GM, Matthews SJ, Hale CB,
                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB20576 standard; protein; 107
    53.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-JAN-2000; 2000WO-GB00254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99GB-0001897
                        Matches 322; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 diarrhoea; antibacterial
                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200045173-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-AUG-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB20576;
       Query Match
                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                               288
                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB20576
                                            ö
                                                           g
                                                                             ò
                                                                                              q
                                                                                                                 ð
                                                                                                                                g
                                                                                                                                                   ö
                                                                                                                                                                    g
                                                                                                                                                                                    δ
                                                                                                                                                                                                       a
                                                                                                                                                                                                                            φ
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                         Q
```

```
The present invention describes a method of screening for an inhibitor of intimin binding to eukaryotic cells. The method comprises exposing an intimin polypeptide having a Tir-independent cell binding activity to test agents, and obtaining an inhibitor based on its ability to bind the polypeptide. The inhibitors are used in the prevention, treatment and/or diagnosis of bacterial infections, preferably by enteropathic and/or enterohaemorrhagic Escherichia coll, Shiga toxigence E. coll, Hafnia chrein or Citrobacter freundii, or especially E. coll 0157:H7. The infections cause a histopathological effect known as attachment and effacement on intestinal epithelial cells. The inhibitors can be used to produce food supplements or additives, especially where the food is a milk substitute. The method can be used to sort cells based on their ability to bind to a Tir independent cell binding domain of an intimin polypeptide. Polypeptides having Tir-independent intimin binding activity can be used to produce a vaccine against a bacterial disease. The present sequence represents a specifically claimed intimin Creminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        255 QALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSGELX 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            for
                                                              Screening for inhibitors of intimin binding to eukaryotic cells, use in diagnosing, preventing and treating bacterial infections, especially Escherichia coli 0157:H7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          translocated intimin receptor N-terminal peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 530; DB 21; Length 107;
Pred. No. 3.1e-34;
0; Mismatches 1; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tir; translocated intimin receptor; Hp90; enteropathogenic;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Stein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EPEC; infection; diagnosis; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kenny B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ā
                                                                                                                                                                                                         Claim 8; Page 76; 96pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY06213 standard; Peptide; 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 18.9%;
Best Local Similarity 99.1%;
Matches 106; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97us-0065130.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-CA01042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Finlay BB,
WPI; 2000-499357/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          107 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        present invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Escherichia coli
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                            ò
SSSSSSSSSSSXX
                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This represents a UspAl antigen of Moraxella catarrhalis strain 035E. Nucleic acid sequences encoding the UspAl and A2 antigens of fm. catarrhalis isolates 035E, 046E, Th2A24 and TTA37 can be used in genetic vaccination. An antigenic composition or vaccine containing antigenic peptides from UspAl or UspA2 antigens are used to induce an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Moraxella catarrhalis; UspAl; UspA2; antigen; genetic vaccination; vaccine; otitis media; sinusitis; lower respiratory tract infection; immunity enhancer; immunoassay reagent.
                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New peptide(s) containing the core epitope of Moraxella catarrhalis Usp proteins - useful in, e.g. vaccines to prevent or treat M. catarrhalis infection, and antibodies for passive immunisation.
                                                        sequence represents the N-terminal sequence of Tir (see
                                                                  also AAY06220), a novel translocated intimin receptor from an enteropathogenic Escherichia coli (EPEC) strain. The 78 kDa EPEC protein is secreted by the bacterial pathogen. Diagnosis of disease caused by pathogenic E. coli can be performed by use of antibodies that bind to Tir to detect the protein, or the use of nucleic acid probes for detection of nucleic acids encoding Tir. Kit for the detection of Tir-producing E. coli is provided. Also protective immune response, and a method for screening for compounds which interfere with the binding of bacterial pathogens
  ρλ
 New translocated intimin receptor useful for treating infection enteropathogenic or enterohemorrhagic Escherichia coli
                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                 Length 30;
                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fredenburg R, Hansen EJ;
                                                                                                                                                                                                                                                   DB 20;
1.9e-06;
                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                            Local Similarity 100.0%; Pred. No. 1.5 nes 30; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 22; Pages 133-136; 237pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   M. catarrhalis strain 035E UspAl antigen
                                                                                                                                                                                                                                                     5.9%; Score 165;
                                                                                                                                                                                                                                                                                                             2 PIGNLGNNVNGNHLIPPAPPLPSQTDGAAR 31
                                                                                                                                                                                                                                                                                                                                                                                Ä
                                      Example 1; Page 37; 91pp; English
                                                                                                                                                                                                                                                                                                                                                                                AAW68201 standard; Protein; 831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cope LD, Fiske MJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97WO-US23930.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   96US-0033598.
                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TEXA ) UNIV TEXAS SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Moraxella catarrhalis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1998-377595/32.
                                                                                                                                                                                              their receptors
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAV41341
                                                                                                                                                                                                                    30 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-DEC-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09828333-A2
                                                                                                                                                                                                                                                                                                                                                                                                                             07-0CT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Aebi C, Co
Maciver I;
                                                                                                                                                                                                                                                                                                                                                                                                        AAW68201;
                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                           ស
                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                        AAW68201
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                        g
```

```
21;
treat infections such as otitis media, sinusitis, lower respiratory tract infections such as otitis media, sinusitis, lower respiratory tract infections. They can also be used as immunity enhancers for other bacterial, parasitic or viral antigens, to raise antibodies and as immunoassay reagents for detecting specific antibodies. The antibodies are useful for passive immunisation and as immunoassay reagents. Detection of the epitopic core sequence, by immunoassay or by FCR, is used to ofagnose infection. The USP antigens encoding nucleic acid sequences are also used to produce recombinant proteins and for screening for potential anti-M. catarrhalis agents, while their fragments are useful as diagnostic probes or primers or to isolate variant sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EE-----SSALIVAGGIGAGVTTAL 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            387 HRRNOPAE----QTTTTTHTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  333 dikdlqkevkgldnevgelsrdinslhdvtdnqqddikelkrgvkeldnevgvlsrdins 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----ASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPE 262
                                                                                                                                                                                                                                                                                                                                                                                                                                         74 SRLAAATSETCLLGGFEVLHDKGPLDILNTQIGP---SAFRVEVQADGTHAAIGEKNGLE 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              226 sgttgvtsnsvilgneta--gkgattvknaevgglsltgfageskaengvvsvgsegger 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    131 VSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVAS----DIAEARTRILAKLD 186
                                                                                                                                                                                                                                                                                                                                                    263 PDDPTTTDPDQAANAAESATKDQLTQEAFKNPE----NQKVNIDAN------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           556 teaidalnkassentgniakngadianninniyelaqqqdqhssdiktlakvsaantdri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----GNAIPSG--ELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              444 DSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSG 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Endoglycoceramidase; activator; glycolipid; enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-----DNHGGRQPKDVDT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW10660 standard; Protein; 580 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endoglycoceramidase activator II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rhodococcus spp. M-777.
                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                           831 AA;
                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Simil
Matches 101; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-JUL-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EP759470-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-FEB-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW10660;
                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW10660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
```

us-09-189-415a-2.rag

Ŋ

Page

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hia; adhesin; Haemophilus influenzae adhesin; NTHi; infection; vaccine; non-typeable Haemophilus influenzae; antiinflammatory; auditory; antibacterial; meningitis; epiglottitis; septicaemia; otitis media; diagnosis; immunogenic; antigen.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SAFRVEVQADGTHAAIGEKNGLEV----SVTLSPQEWSSLQSIDTEGKNRFVFTGG 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RGGSGHPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSET 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --gpdh----dayrdqpvnltakvepgvsggtvafevd--gtpvgtadvmddgaavlp 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALAL-----TPEPDDPTTTDP 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                htfttngthr-----vlarysg--aegispsvslgypvsvteapaadvattitv 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  215 gqfgndanfltflprasapivgtvwaptqcsprdtaa---gplnagagplatiqi---l 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RNSMADSVDSRDIPGLPTNPSR--LAAATSETCLLGGFEVLHD----KGPLDILNTQIGP 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---APPLPSQTDGAARGGTGHLISSTGALGSRSLFSPL 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Haemophilus influenzae adhesin (Hia) protein from type c strain API
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              asltgsldtdgdyayeakflgaeffnpssaaktvtvtsgdigtttsvt------
                                                                                                                                                                                      DNA encoding endo:glyco:cer:amidase\ activator\ polypeptide - for prodn. of recombinant polypeptide, useful in glyco:lipid\ analysis
                                                                                                                                                                                                                                                               DNA encoding endoglycoceramidase activator polypeptide can be use for the prodn. of the recombinant polypeptide, which can be used to research the intracellular function of glycolipids, because it shifts the optimm pH for endoglycoceramidase II towards neutral, allowing the latter to be used to hydrolyse glycolipids
                                                                                                                                                                                                                                                                                                                                                                                                                                                       96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DQAANAAESATKDQLTQEAFKNPENQK------VNIDANGNA 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | |: |: : :| | :| : : dpiastakgs---pvtltarldpadargtvqfklgdvllggpvrvdangva 516
                                                                                                                                                                                                                                                                                                                                                                                                                           Length 580;
                                                                                                                                                                                                                                                                                                                                                                                                                       ; Score 141; DB 18; Length 5; Pred. No. 0.0085; 45; Mismatches 126; Indels
                                                                                                           Ξ
                                                                                                             Sano
                                                                                                           Υ,
                                                                                                             Kurome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB23860 standard; Protein; 2411 AA
                                                                                                                                                                                                                                   Claim 1; Page 13-15; 30pp; English
                                                                                                           Kato I,
                                                                                                                                                                                                                                                                                                                                                                                                                        5.0%;
                                               95JP-0188466.
                96EP-0110513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       84; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNLGNNVNGNHLIPP----
                                                                                                           Izu H, Izumi Y,
                                                                            (TAKI ) TAKARA SHUZO CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Haemophilus influenzae
                                                                                                                                       WPI; 1997-147519/14.
N-PSDB; AAT61050.
                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                              580 AA;
                                                                                                                                                                                                                                                                                                                                             even at pH 7.5
                28-JUN-1996;
                                             29-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-JAN-2001
                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB23860
                                                                                                           Ito M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB23860
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
The present sequence represents a Haemophilus influenzae adhesin

(Hia) protein from the type c Haemophilus influenzae strain API.

(Hia) protein from the type c Haemophilus influenzae strain API.

(Hia) genes and proteins have antiinflammatory, auditory and antibacterial activities, and can be used in the production of a vaccine. An immunogenic composition comprising an Hia gene, a polypeptide encoded by an Hia gene, or a recombinant Hia polypeptide is useful for inducing correction against disease caused by Haemophilus strains in a susceptible host, preferably a human. An Hia protein is useful as an antigen, in immunogenic preparations including vaccines, as a carrier for other lammonogens, and in the generation of diagnostic reagents. Hia is useful for treating diseases caused by the infection of Haemophilus influenzae such as meningitis, epiglottitis, septicaemia and otitis media. Recombinant production of Hia favours high recovery of the low recovery of mative protein from Haemophilus influenzae species. A truncated protein has a significantly higher amount of recovery than a full-length protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    el nucleic acid encoding Hemophilus influenzae adhesin protein, for as antigens and vaccines and for treating Hemophilus influenzae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1161 lggtangrndtgtvinkdgltitlangaaagtdasngntisvtkdgisagnkeitnvksa 1220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1221 lktykdtgntag-atgp---aantaeva-----kgdlvdltkpatgaagngadaka 1267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1328 khtvti--dvaeakvgdglekdtdgkiklkvdntdgnnlltvda-tkgasvakgefnavt 1384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : | : | : | | | : | | : | | atvsdkl:-----slgtngnkvnitsdtkglnfakdsktgddanihl-ngiastltd 1544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MADSYDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQA 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -DGTHAAIGEKNGL-----EVSVTLSPQEW----SSLQSIDTEGKNRFVFTGGRGGSG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               165 HPMVTVASDIAEAR-----TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVV 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SETHTST-TNSSVR-----SDPKFWVSVGAIAAGLAGLAATGIAQALALT 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       261 PEPDDPTTTDPDQAANAAESATK--DQLTQEAFKNPENQKVNID-----ANGNAIPS 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSS 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 LGNNVNGNH----LIPPAPPLPSQTDGAARG---GTGHLISST--GALGSRSLFSPLRNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.0%; Score 139.5; DB 21; Length 2411;
21.5%; Pred. No. 0.079;
ive 83; Mismatches 228; Indels 145;
                                                                                                                                                                                                                                                                                                                                                                                  Klein MH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1; Fig 24; 275pp; English.
                                                                                                                                                     2000WO-CA00289
                                                                                                                                                                                                                                99US-0268347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                  Loosmore SM, Yang Y,
                                                                                                                                                                                                                                                                                                       (CONN-) CONNAUGHT LAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-618897/59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2411 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAA92499
WO200055191-A2
                                                                                                                                                     16-MAR-2000;
                                                                                                                                                                                                                                16-MAR-1999;
                                                                          21-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             311
       $\circ{\circ}{\circ}\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\cir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
```

28;

9

```
1244 av--knanevefv-----
                                                                                                                                                                                                                                                                                                                                                                                                                        19-APR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-JAN-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gross HJ,
                                                                                                                                                                                                                                                                                                                                                                                                    AAB46770;
                                                                                                                                                                328
                                                                          230
                                                                                                                                                                                                                                                                                                                                                                  AAB46770
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                      δŏ
                                                                                                                                          g
                                                                                                                                                               à
                                                                                                                                                                                     pp
                                                                                                                                                                                                            δ
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                      g
                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1199 -----gntisvtkdgisagnkeitnvksalktykdtgnt-----adetgdkefha 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1163 lggtang------rndtgtvinkdg-----ltitlangaaagtdasn 1198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPL----DILNTQIGPSAFRVEVQADGTHA 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 AIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR--- 178
                                                           -----gg----ngitdnekkraasv 1571
371 ALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDR 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels 171; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 LGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRD 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Haemophilus adhesion protein HA2 (AAR99393) is associated with the formation of surface fibrils involved in adhesion to various host cells; it is also referred to hsf (Haemophilus surface fibrils). Its amino acid sequence was deduced from a genomic DNA clone (AAF41476) derived from Haemophilus influenzae type b strain C65. Large quantities of recombinant HA2 can be produced in transformed prokaryotic or eukaryotic host cells, for use in vaccines against
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 use in
                                            431 RDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.9%; Score 138.5; DB 17; Length 2353; 22.0%; Pred. No. 0.091;
Live 62; Mismatches 207; Indels 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 - for
                                                                                                                                                                                                                                                         Haemophilus adhesion protein; HA2; hsf protein; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Recombinant Haemophilus adhesion proteins HA1 and HA2 vaccines against H. influenzae infection.
                                                                                                     | ::||:||
1611 dkdttsvtveskdngkrtevkigaktsvikdhngklftgke 1651
                                                                                        491 SGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGE 531
                                                                                                                                                                                                                                                                                Haemophilus influenzae type b strain C54,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Page 66-73; 120pp; English.
                                                                                                                                                                      Ā
                                                                                                                                                                                                                                      Haemophilus adhesion protein HA2
                                                                                                                                                                      AAR99393 standard; Protein; 2353
                                                                                                                                                                                                                                                                                                                                                                        95US-0409995.
                                                                                                                                                                                                                                                                                                                                                   96WO-US04031
                                                                                                                                                                                                                                                                                                                                                                                                                               St Geme JW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 124; Conservative
                          1545 tl----lnsgattnl----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   H. influenzae infection
                                                                                                                                                                                                                                                                                                                                                                                              (UYSL-) UNIV ST LOUIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1996-455364/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2353 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAT41476
                                                                                                                                                                                                                                                                                                                                                                                                                               Barenkamp SJ,
                                                                                                                                                                                                                                                                                                         WO9630519-A1
                                                                                                                                                                                                                                                                                                                                                   22-MAR-1996;
                                                                                                                                                                                                                                                                                                                                                                        24-MAR-1995;
                                                                                                                                                                                                                                                                                                                              03-OCT-1996
                                                                                                                                                                                                                 15-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 5;
                                                                                                                                                                                            AAR99393;
                                                                                                                                                                                                                                                                                                                                                                                                          ( MIND)
                                                                                                                                                            AAR99393
                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à.
                                                                   q
                         g
                                                à
                                                                                            ò
```

```
This invention describes a novel nucleic acid (I) which has antitumor activity. (I) encode proteins that interact with the Fanconi anemia protein of complementation group A, so may be part of the complex or signal transduction cascade which causes the defect responsible for Fanconi anemia. Polypeptides (II) encoded by (I) are used as immunogens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       proteins,
                                                                                                                                                                                                                    1445 lgtngnkvnitsdtkglnfakdsktgddanihl-ngiastltdtl-----lnsgattnl- 1497
|||| : :| ||: |:|||:
----gkngatvsaktdnngkhtvti--dvaeakvgd 1284
                                                                                                              388 RRNQPAEQTTTTTTHTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSS 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        448 EVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQ 507
                                                                                                                                                                                                                                                                                           278 AESATK--DQLTQEAFKNPENQKVNID-----ANGNAIPSGELXDDIVEQIAQQAKE 327
                                                                                                                                                                                                                                                                                                                                                                                                      AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FANCIP2; FANCIP3; antitumor; signal transduction cascade; immunogen; Fanconi anemia complementation group A protein; DNA-repair defect; cell-cycle disorder; cytopenia; tumorgenesis; tumor progression;
                                                                                                                                                                                                                                                                                                                               ----TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTST-TNSSVR---
                                                                                                                                                                                        -----SDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New nucleic acid encoding Fanconi anemia protein interacting luseful e.g. for diagnosis and treatment of DNA-repair defects
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wagner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----gg----ngitdnekkraasvkd-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Herterich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FANCIP3 protein fragment SEQ ID NO 10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB46770 standard; Protein; 3096 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hanenberg H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 1; Page 49-58; 60pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        508 STYALLANSGGLRLGMGGLTSGGE 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (MULT-) MULTIGENE BIOTECH GMBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-JUN-2000; 2000WO-EP05878.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Reuter T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-102892/11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200100822-A2
```

```
125; Conservative
                                                          venezuelae.
                                                                                                                                                                                              (MINU ) UNIV MINNESOTA.
  Polyketide synthase; p
polyhydroxybutyrate; b
metabolic engineering.
                                                                                                                                                                                                                                                     WPI; 1997-341701/31.
N-PSDB; AAT68715.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4630 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                          industrial use
                                                                                    WO9722711-A1
                                                                                                                                           18-DEC-1996;
                                                                                                                                                                    19-DEC-1995;
                                                                                                                                                                                                                         Sherman DH,
                                                                                                               26-JUN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local
Matches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                227
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                    33;
for preparation of antibodies (Ab), also for identifying specific modulators (A). Any of (I) (or vectors or cells that contain (I)), (II), Ab and (A) are useful for: (a) diagnosis of diseases (or predisposition to them) associated with DNA-repair defects, cell-cycle disorders, cytopenia, tumorgenesis and tumor trogression; and (b) for treatment (including by gene therapy) or prevention of these diseases.
                                                                                                                                                                                                                        1916 pasplys -- glalagl -- -- -- nlelysgl - nglasnlysmetpheglnasple 1960
                                                                                                                                                                                                                                                                                                                                     2021 snlysleleaspalatyrthrglnileserasplelysargmetileserlysleglala 2080
                                                                                                                                                                                                                                                                                                                                                                                           2081 glnvallysglnvalglhisglasnmetleserlearghisasnserargilehisvala 2140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2244 s--erasnvalaspasnglnle--glglthrcysserleglyhisargserprleglly 2298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1961 leglygltyrglserleglylysglhisargargvallysaspalaleasnthrthrqla 2020
                                                                                                                                                                                                                                                                                                                                                                                                                                               2141 rgpr--serargalaa----snthrlealathrserasp---valserargargly 2187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    strpleileprglyalagityrserilephethrglyglnprleaspt----hrglnasp 2243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2299 saspserserprglyserserserthrserleleilelyslysglnarggl-----th 2351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2407 ntrpglythrglnthrgllysglaspthrserasnileasnp----rargglnthrglt 2461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2462 hrservalasnalaserargserprg11-----yscysalaglng1n-argglnlys 2512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 429 DRRDS--------QSVASTHWSDSSSEVV---NPYAEVGGARNSLSAHQPEE 470
                                                                                                                                                                                                                                                                                                           PSAFRVEVQADGTHAAIGEKNGLEV--SVTLSPQEWSSLQSIDT------E 149
                                                                                                                                                                                                                                                                                                                                                                                                                       RQPKDVDTRSVGVGSASGIDDGVVSETH---TSTTNSSVRSDPKFWVSVGAIAAGLAGLA 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            250 AT-----GIAQALALT-----PEPDD-------PTTTDPDQAANAAE 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 280 SATKDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVES 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RRNOPAEQTITITHTVVQQQTG------GIPQHKVALMPQERRRFS 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       471 HIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGL---- 526
                                                                                                                                                                                                                                                     LAAATSETC-----LLG------IGFEVLHDKGP---LDILNTQ----IG 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      340 NAQAQQRYE------DQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
                                                                                                                                                                      Gaps
                                                                                                                                                                                                PAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIP--GLPTNPSR 75
                                                                                                                                                                                                                                                                                                                                                                GKNRFVFTGGR------GGSGHPMVTVASDIAEARTRILAKLDPDNHG------G
                                                                                                                                           Length 3096;
                                                                                                                                                                    Indels 219;
                                                                                                                                           DB 22;
                                                                                                                                         ; Score 136.5; DB 22;
; Pred. No. 0.19;
78; Mismatches 245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Streptomyces venezuelae polyketide synthase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2513 argleasnseralaserglnargser 2538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            527 ----TSGGETAVSSVNAAPTPGPVR 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW19629 standard; Protein; 4630 AA
                                                                                                                                         4.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                    Conservative
                                                                                                                                                      Similarity
                                                                                                3096 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-SEP-1997
                                                                                                                                                       Best Local Sim
Matches 144;
                                                                                                  Sequence
                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW19629
                                                                                                                                                                                                                                                                                                           107
                                                                                                                                                                                                                                                                                                                                                                  150
                                                                                                                                                                                                                                                                                                                                                                                                                     193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2188
                                                                                                                                                                                              18
                                                                                                                                                                                                                                                     92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 1
    888888888
                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                              Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .
d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          XEXEXEX
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
```

```
1215 ggesslalaggatvmstpgmfvefsrqrglaadgrskafsdsadgtswaegv--gllvve 1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             grvirgalad - agltpadvdaveahgtgtrlg - dpieaeailgtygrdrgegaplqlg 1366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1422 er----ewpetgr---prraaissfgisgtnahivveqapeage----- 1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1156 prmhdg-aegveghlltgttpsvmsgriayqlgltgpavtvdtacssslvalhlavrslr 1214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       slksn----iahaqaaagvgglikmvlamrhgvlprtlhvdrpttr-vdweaggvellte 1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AATSETCLLGGFEVLHDKGPLDILNTQIGPSA----FRVEVQADGTHAAIGEKNGLEVSV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134 TLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIA-------EAR 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAE--QTTTTTT 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 Polypeptide sequences (AAW19629-30 and AAW00918) can be deduced frot the vep ORF1 polyketide synthase (PKS) gene cluster (AAT68715) of Streptomyces venezuelae. The sequence data indicate that the PKS sequence data indicate that the PKS gene cluster encodes a polyene of 12 carbons. The vep gene cluster contains 5 PKS modules, plus a 5' loading module and a 3' end domain. Each of the sequenced modules includes a keto-ACP, an acyltransferase, a dehydratase, a keto-reductase and an acyl carrier protein domain. A novel expression cassette encoding the first module from the vep gene cluster and module 7 from the Streptomyces tylp gene cluster has polyhydroxyalkanoate (PHA) monomer synthase activity and can be used for PHA prodn. in host (esp. insect) cells for use as a biodegradable polymer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23 PSQTDGAARGGTGHLISST--GALGSRSLFSPLRNSMADSVD---SRDIPGLPTNPSRLA 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEP----DDPTTTDPDQAANAAESATK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Expression cassettes for production of polyhydroxyalkanoate(s) provide wide range of biodegradable polymers for medical or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 4630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187;
synthase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              179 TRILAKLDPDNHGGRQPKDVD-TRSVGVGSASGIDDGVVSETHTSTTN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
polyhydroxyalkanoate monomer sylbiodegradable polymer; vep gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.9%; Score 136.5; DB 18;
22.2%; Pred. No. 0.33;
ve 49; Mismatches 203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 55; Fig 23; 91pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                    96WO-US20119.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95US-0008847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Williams MD,
```

œ

```
vulnearay; artipsoriatic; antiparkinsonian; procession neuroprotective; antiporatary; antipsoriatic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coaquiant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressaye; antidifammatory; antianaemic; gene therapy; cancer; proliferative disorder; hypotensive; antianaemic; gene therapy; cancer; proliferative disorder; hypoterension; neurodegenerative disorder; osteoarthritis; apraft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; astever combined immunodeficiency; malaria; autoimmune disorder; asthma; allery; aplastic anaemia; nocturnal heemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coaquiation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticarkinsonian; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive;
---aavtttapeageageaadttattt 1482
                                                                                                                                               1521 ---pdvtvadlgrslaartafehkaalttatrdella--gldal----grgeqatglvtg 1571
                                    402 HTVVQQQTGGIPQ-----HKVALMPQ--ERRRFSDRRDSQGSVASTHWSDSSSE 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    open reading frame; ORFX; detection; cytostatic; hepatotropic;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease
                                                                                                              449 VVNPYAEVGGARNSLSAHQPEEH-----IYDEVAADPGYSVIQNFSGSG-PVTGRLIG
                                                                                                                                                                                                                                                                                                                                                                                                                                             Human ORFX ORF709 polypeptide sequence SEQ ID NO:1418.
                                                                          1483 paav----gvpepvrapvvvsardaaalraqavrlrtfldgr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 11; Page 1203-1206; 5507pp; English
                                                                                                                                                                                                                                                                                                                                AAB40945 standard; Protein; 1532 AA
                                                                                                                                                                                                                                  1572 eparaartaf-lftgggagrvamg 1594
                                                                                                                                                                                             501 TPGGGIQSTYALLANSGGLRLGMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0127636.
99US-0127728.
2000US-0540763.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-MAR-2000; 2000WO-US08621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0127607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          thrombosis; contraceptive
                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Leach M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CURA-) CURAGEN CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-602362/57.
N-PSDB; AAC75154.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200058473-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Shimkets RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      05-APR-1999;
30-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                             08-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                       AAB40945;
        1459
                                                                                                                                                                                                                                                                                                          AAB40945
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                       à
                                                                                                                                                           ద
                                                                                                                                                                                             à
                                                                                                                                                                                                                                  g
```

```
antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating the presence of or predisposition to, or preventing or treating the protein associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy rections. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allastic anaemia, burns, wounds, bone and cartilage damage, necturnal haemoglobinuria, antilniammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.
                                                                                                                                                                                                                                                                                                                                                                                                                         28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        961 -----tlspsgsgktfttalisnatplpvtyassastghttplhvtdassvstghat 1012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 490 GSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGET-----AVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFS 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      372 LIVAGGIGAGVITALHRRNQPAEQTITITITITIVVQQQIGGIPQHKVAL--MPQERRRFSD 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          510 lpststgaatrlvtgn--pstgaagt----iprvpskvsaige---pgepttysshsttl 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         pkttgagaqtqwtqetgttg--eallsspsysvtqmiktatspssspmldrhtsqqitta 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       157
                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                LPSQTDGAA-RGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTN----- 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                813 sgasgttpsgsegistsgettr---fssnps----rdshttgsttellsasashg---a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rgsdtislas-qatdtfstv--pptppsitssgltspqtqth------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   862 ipvstgmassivpgtfhptlseastagrptggssptspsaspgetaaisrmagtgrtrts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73 PSRLAAA----TSETCLLGGFEVLHDKGPLDI---LNTQIGPSAF-----RVEVQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116 ADGTHAAIGEKNGLEVSVTLSPQEWSSL------OSIDTEGKNRFVFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               619 pstnhstih----ststspqespavsqrghtqapqttqesqttrsvspmtdtktvtt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         158 GGRG--GSGH------PMVTVASDIAEARTRILAKLDPDNHGGRQPKDV----DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                672 pgssftasghspselvpqdaptisaattfapapt-----gdghttqapttalqatpss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          201 RSVGVGSASGIDDGVVSETHTSTTNSSVRSDP----KFWVSVGAIAAGLAGLAATGIAQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         257 LALTPEPDDPTTTDPDQAA----NAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        312 ELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSA
                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 178;
                                                                                                                                                                                                                                                                                                                                                                                 Length 1532;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61; Mismatches 237;
                                                                                                                                                                                                                                                                                                                                                                                   4.8%; Score 135; DB 21; 21.1%; Pred. No. 0.095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY27230 standard; Protein; 1978 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 127; Conservative
                                                                                                                                                                                                                                                                                                                              1532 AA;
                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1013 plp 1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             541 PTP 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY27230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY27230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
```

```
1401 qvqlay-dkwdykqegltragaaivtiivtaltygygata----aggvaasgsstaaaag 1455
                                                              taatttaaattvstatamqtaalaslysqaavsii-----nnkgdvgkalkdlgtsdt 1508
                                                                                                                        1509 vkqivts-altagalnqmgadiaqlnskvrtelfsstgnqtianlg-----grlatnls 1561
                                 NQPAEQTTTTTTTTTTTTTTVQQQTGGI-----PQHKVALMPQERRRFSDRRDSQGSVASTHWSDS 445
                                                                                            446 SSEVVNPYAEVGGARNSLSAH--QPEEHIYDEVAADPGYSVIQNFSGSGPVTGRL-IGTP 502
                                                                                                                                                                                                                                                                                                                                                                                        S-layer; rsaA gene; Caulobacter; vaccine; antigenic; ligand; enzyme;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New DNA containing sequence for C-terminal region of Caulobacter S-layer protein - expressed as fusion proteins containing antigenic peptides in Caulobacter, useful as live vaccines
                                                                                                                                                                                                                                                                                                                                                                                                    metallothionein; heavy metal; water; sewage; xylanase; cellulase;
                                                                                                                                                                                   1562 nagisagintavnggslkdnlgnaalgalvnsfqgeaa 1599
                                                                                                                                                         GQGIQSTYALLANSGGLRLGMGGLTSG-----GETA 533
                                                                                                                                                                                                                                                                                                                                                         Caulobacter crescentus S-layer rsaA protein.
                                                                                                                                                                                                                                                                   Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Smit J;
                                                                                                                                                                                                                                                                 AAW37490 standard; Protein; 1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 2; Fig 6; 58pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BRITISH COLUMBIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96US-0614377.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97WO-CA00167
                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bingle WH, Nomellini JF,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caulobacter crescentus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1997-470880/43.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAV01866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1026
                                                                                                                                                                                                                                                                                                                                                                                                                        wood pulping.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9734000-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (UYBR-) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-MAR-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12-MAR-1996;
                                                                                                                                                                                                                                                                                                                              20-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-SEP-1997
                                                                                                                                                                                                                                                                                                AAW37490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                 390
                                                                                                                                                         503
                                                                1456
                                                                                                                                                                                                                                                    AAW37490
                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                요
                                                                                                                        g
                                                                                                                                                        δ
                              δ
                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            meningitidis (strains A and B) and nucleic acid sequences (AAX99123-167) encoding the proteins. Compositions comprising the protein, nucleic acid or antibody specific to the protein are useful as pharmaceuticals, e.g. a vaccine composition or a diagnostic composition. The composition is also useful for treating or preventing an infection due to Neisserial bacteria, especially Neisseria meningitidis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1293 a----digagygekaradakiilkgivnrigseekletnstvwgkg---agrgstietlk 1345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1072 lidgitdqyeigkptykshydkaalnkpsrltgrtgvsihaaaalddariiigaseikap 1131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | :||
sgsidikahsdivleagqndaytflktkgksgkiirktkftstrdhlimpapveltangi 1191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1235 -dvgksrrfigikvgksnysknelnetklpvrvvaqtaatrsgwdt-vlegtefkttlag 1292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----VEVQADGT 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120 HAAIGEK-----NGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVA 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                228 VRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLT 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           288 QEAFKNPENQKVNIDANGN---AIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNA--- 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QAQQRYEDQHARRQEELQLSSG------IGYGLSSALIVAGGIGA-GVTTALHRR 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   172 SDIAEARTRILAKLDPDNHGGRQPKD--VDTRSVGVGSA--SGIDDGVVSETHTSTTNSS 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----AAATSETCLLGGFEVLHD 94
                                                                protein; pharmaceutical; vaccine; diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                      New protein and its nucleotide sequence, useful in vaccines or diagnostic compositions for treating and/or preventing Neisseria meningitidis infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 20; Length 1978;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention provides proteins (AAY27201-245) from Neisseria
                                                                                                                                                                                                                                                                                                                            Scarlato V;
                                Amino acid sequence of N. meningitidis protein ORF114-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 69; Mismatches 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----LNTQ----IGPSAFR---
                                                                                                                                                                                                                                                                                                                              Rappuoli R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4.8%; Score 133.5; DE 22.1%; Pred. No. 0.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----NPSRL-----
                                                                                                                                                                                                                                                                                                                              Pizza M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 4; Page 88; 123pp; English.
                                                            Neisseria meningitidis protein;
bacterial infection; treatment.
                                                                                                                                                                                                                                                 98GB-0000760.
98GB-0019015.
                                                                                                                                                                                                                                 98GB-0022143
                                                                                                                                                                                                    99WO-IB00103
24-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57 MADSVDSRDIPGLPT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                            Grand1 G, Masignani V,
                                                                                                         Neisseria meningitidis
                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-444400/37.
N-PSDB; AAX99152.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1978 AA;
                                                                                                                                                                                                                                                                                                (CHIR-) CHIRON SPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             95 KGPLDI ---
                                                                                                                                        WO9936544-A2
                                                                                                                                                                                                    14-JAN-1999;
                                                                                                                                                                                                                                                                 01-SEP-1998;
                                                                                                                                                                                                                                 09-OCT-1998;
                                                                                                                                                                      22-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        342
```

ö 셤 ŏ g ò g ά g ò g ολ g

The present sequence represents a Caulobacter S-layer protein used in an example of the present invention. A new DNA construct has been developed which contains at least one restriction site for insertion of DNA upstream of DNA encoding a C-terminal region of at least the last 82 amino acids (aa) of Caulobacter S-layer protein. Caulobacter containing the DNA constructs above additionally containing a sequence encoding a heterologous polypeptide, are particularly useful in live vaccines (where the heterologous polypeptide is an antigen). They can also be used for production of e.g. ligands, enzymes or other proteins, e.g. metallothioneins to remove heavy metals from water or sewage, or xylanase or cellulase for use in wood pulping. All known Caulobacter strains are harmless, and stable in outdoor environments, including water (so suitable for vaccinating fish) or soil. They are well suited for growing in biofilm reactors and produce S-layer proteins, which is an ideal system for presentation of antigens, at high level. Score 132.5; DB 18; Length 1026; Pred. No. 0.086; 4.7%; 21.8%;

```
Sequence
                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                               Local
         (UYBR-)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ΟŸ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers 692..693
/note= "Asp-Pro dipeptide present in S-layer secretion signal sequence. It is a site where a fusion protein comprising a target protein and the secretion signal is cleaved"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Surface layer protein; S-layer secretion signal; antibiotic; vaccine; recombinant fusion protein cleavage; enzyme; protein polymer; antibacterial enzyme; foodstuff.
                                                                                                                                                                                                                                                                               360 LSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVAL 419
                                                                                                                                                                                                                                                                                                   --sgtgtsl-----gigrgalta----tptantltlnvnglt--ttgaitdseaa- 560
                                                                                                                                                                                                                                                                                                                                                                                 || | :| : |: || || || || || aaaltgitvasyatlgaelatg-lvftggagrdsillgattkalvmgagddtvtvssa 663
                                                                                                                                                                                                                                     PSGELXD--DIVEQIAQQAKEAGEVAR-----QQAVESNAQAQQRYEDQHARRQEELQ 359
                                                                                   281 ltvgdtlsggagtdvlnwvqaaavtalptgvtisgietmnvtsgaaitlntssgvtglta 340
                                                                                                                                                                       424
                                        230 tdnaagvnlftaypssgvsgstl-----slttgtdt--ltgtanndtfvagevagaat 280
                                                               LGGFEVLHDKGPLDILN------TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT 134
                     TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85
                                                                                                                                                                                                                                                                                                                                       VAADPGYSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GL
                                                                                                                                                                    249 AATGIAQALALIPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAI
                                                                                                                                                                                                                                                                                                                          MPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDE
                                                                                                           LSPQEWSSLQSIDT-EGKNRFVFTGGRGGS-----GHPMVTVASDIAEARTRILAKLDPD
                                                                                                                             341 Intntsgaaqtvtagagqnltattaaqaannvavdgranvtvas-----tgvts----
                                                                                                                                                    NHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGL
                                                                                                                                                                                                                   tavtvaqtag---navnttltqadvtvtgnssttavtvtqta------aatagat
                                                                                                                                                                                                                                                    :| : | | || ||::|
47] vagrvngavtitdsaaasattagkiatvtlgsfgaatidssalttvnl-------
 155;
 64; Mismatches 222; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Caulobacter crescentus surface layer protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY44757 standard; Protein; 1026 AA
                                                                                                                                                                                                                                                                                                                                                                                                                 RLGMGGLTSGG---ETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                   664 tlgaggsvnggdgtdvlvanvngs 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99WO-CA00637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98CA-2237704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04-MAY-2000 (first entry)
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caulobacter crescentus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200004170-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cleavage-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY44757;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
                                                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                                                                       476
                                                                                                                                                                                                                                                                                                                                                                                                                 520
                                                                                                                                                                                                                                           309
                                                                                                                                                                                                                                                                                                        519
                                                                                                                                                                                                                                                                                                                                                  561
     Matches
                                                                   86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                    ρp
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                            ò
                                                                                       q
                                                                                                             ò
                                                                                                                               g
                                                                                                                                                        ò
                                                                                                                                                                       g
                                                                                                                                                                                           à
                                                                                                                                                                                                                                          ă
```

```
The patent discloses a method for cleaving a recombinant fusion protein which is produced by Caulobacter and consists of Caulobacter surface layer (S-layer) protein (containing the C-terminal secretion signal) and a target protein heterologous to Caulobacter. The cleavage of target protein from the S-layer protein is carried out under mild acid protein from the S-layer protein is carried out under mild acid protein so that cleavage occurs at aspartate conditions so that cleavage occurs at aspartate. The cleavage is accomplished while the fusion protein is in The cleavage is accomplished while the fusion protein is in a insoluble aggregate form which facilitates purification of the protein. The method is useful for producing pure proteins including recombinant human and animal therapeutic antibiotic and vaccine peptides, enzymes, protein polymers, and antibacterial enzymes for foodstuffs. The present sequence is a S-layer protein from C crescentus.

The present secretion signal, corresponding to the C-terminal portion of the protein from amino acid 690 onwards, is fused with a target sequence the protein from amino acid 690 onwards, is fused with a target sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27;
                                                                                                                                                                                                                                                                                                                                                                                                                       discloses a method for cleaving a recombinant fusion protein
                                                                                                                                                                                                                                      Cleavage of Caulobacter produced recombinant fusion proteins useful for producing vaccine peptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --sgtgtsl----gigrgalta----tptantltlnvnglt--ttgaitdseaa- 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            || | :| : || || || || || || aaaltgitvanyagddtvtvssa 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tavtvagtag---navnttltgadvtvtgnssttavtvtgta------aatagat 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     249 AATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             189 NHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTINSSVRSDPKFWVSVGAIAAGLAGL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26 TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PSGELXD--DIVEQIAQQAKEAGEVAR-----QQAVESNAQAQQRYEDQHARRQEELQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------addgfttiniagstasstiaslvaadattlnisgdarvtitsh-----t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VAADPGYSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     135 LSPQEWSSLQSIDT-EGKNRFVFTGGRGGS-----GHPMVTVASDIAEARTRILAKLDPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      vagrvngavtitdsaaasattagkiatvtlgsfgaatidssalttvnl-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              230 tdnaagvnlftaypssgysgstl-----slttgtdt--ltgtanndtfvagevagaat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             86 LGGFEVLHDKGPLDILN-----TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              281 ltvgdtlsggagtdvlnwvqaaavtalptgvtisgietmnvtsgaaitlntssgvtglta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              222; Indels 155;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.7%; Score 132.5; DB 21.8%; Pred. No. 0.086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64; Mismatches
                                                                                                                                                                                                                                                                                                                                                              Example 1; Pages 21-23; 33pp; English
BRITISH COLUMBIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                           WPI; 2000-182434/16.
N-PSDB; AAZ50079.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1026 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
   UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 123;
```

4630 AA;

Sequence synthase

pikromycin;

SXC

```
The inversion betacase to an instance and purities nature actual sequence of the inversion to the inversion betacase to an instance of the decision of the eryc gene cluster of Saccharopolyspora erythraea or Streptomyces antibioticus. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide biosynthetic gene cluster encodes proteins which synthesise methymycin, parabomycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic gene clusters are useful for synthesis of methymycin, plkromycin, plkromycin, and narbomycin. The alternative termination of polyketide proteins are useful for synthesis of methymycin, plkromycin, and narbomycin. The alternative termination of polyketide synthesis may be useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA) monomers. The compounds produced by the recombinant host cells are useful as biopolymers, e.g., in packaging or biomedical applications, to engineer PHA monomers synthases or to prepare to biologically active agents to treat asthma, chronic obstructive pulmonary diseases as well as other diseases involving respiratory inflammation, cholesterol-lowering agents or macrolide—based antibiotics which are active against a variety of organisms, e.g., bacteria, including as uniti-dury respiratory pathogens, as well as a particular or the active against a variety of organisms, e.g., bacteria, including as uniti-dury respiratory pathogens, as well as a particular parasition and antibotics.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     as viral parasitic pathogens, or as crop protection agents (e.g., fungicides or insecticides) via expression of polyketides in plants. Sequences AAY77181-Y77189 and AAY77199 represent Streptomyces venezuelae ATCC 15439 protein sequences. These are encoded by a DNA sequence designated vep ORF 1, which actually contains 3 open reading frames. The vep ORF 1 protein is defined in the specification as a PHA monomer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           invention relates to an isolated and purified nucleic acid segment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Desosamine and macrolide biosynthetic gene clusters, useful for, e .g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase; blopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma, chronic obstructive pulmonary disease; respiratory inflammation; hypercholesterolaemia; crop protection agent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Desosamine biosynthesis; macrolide; polyketide; methymycin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S. venezuelae vep ORF 1 amino acid sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  synthesis of methymycin and pikromycin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 3; Figure 23; 438pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zhao L;
                                                                                                                                                                                                                                                                        AAY77177 standard; Protein; 4630 AA
520 RLGMGGLTSGG---ETAVSSVNAA 540
                                                 || || :|| :|| : ||::|| : 664 tlgaggsvnggdgtdvlvanvngs 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Streptomyces venezuelae ATCC15439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-US14398.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Xue Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0105537
                                                                                                                                                                                                                                                                                                                                                                                                                                 05-JUN-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MINU ) UNIV MINNESOTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Liu H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000-160679/14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAZ87283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200000620-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sherman DH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000
                                                                                                                                                                                                                                                                                                                                                   AAY77177;
                                                                                                                                                                                                  RESULT 1

TO AAY7

XX XX AAY7

DE XX AAY7

DE S. v AAY7

NAME OF S. v 
ò
                                                                          g
```

```
27;
                                                                                         -------rnghpvlavirgsavnqdgasngltapngpsq 1310
                                                                                                                                                                                                                                                                                                                      :: :|| || : || | : || || || er-----ewpetgr----prraaissfgisgtnahivvegapeage------ 1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----aavtttapeageaadttatt 1482
                                                                                                                                                                                                                                                                                                                                                                                                          1367 slksn----ighaqaaagygglikmvlamrhgvlprtlhvdrpttr-vdweaggvellte 1421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      402 HTVVQQQTGGIPQ-----HKVALMPQ--ERRRFSDRRDSQGSVASTHWSDSSSE 448
4.7%; Score 132.5; DB 21; Length 4630;
22.2%; Pred. No. 0.68;
Live 48; Mismatches 204; Indels 187; Gaps
                                                                                                                                              78 AATSETCLLGGFEVLHDKGPLDILNTQIGPSA----FRVEVQADGTHAAIGEKNGLEVSV 133
                                                                                                                                                                                                                     ----EAR 178
                                                                                                                                                                                                                                                                                             179 TRILAKLDPDNHGGRQPKDVD-TRSVGVGSASGIDDGVVSETHTSTTN-----S 226
                                                                                                                                                                                                                                                                                                                                                                       SVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEP---DDPTTTDPDQAANAAESATK 283
                                                                                                                                                                                                                                                                                                                                                                                                                                              DQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               449 VVNPYAEVGGARNSLSAHQPEEH-----IYDEVAADPGYSVIQNFSGSG-PVTGRLIG 500
                                                                        23 PSQTDGAARGGTGHLISST--GALGSRSLFSPLRNSMADSVD---SRDIPGLPTNPSRLA 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    344 QQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAE--QTTTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           paav----gvpepvrapvvvsardaaalraqavrlrtfldgr
                                                                                                                                                                                                                     134 TLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1572 eparagrtaf-lftgqgaqrvamg 1594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      501 TPGQGIQSTYALLANSGGLRLGMG 524
                                      Matches 125; Conservative
                    Similarity
                                                                                                                                                                                                                                                         1273 rlsdae-----
                   Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1521
                                                                                                                                                                                                                                                                                                                                                                       227
                                                                                                          g
                                                                                                                                                                                 q
                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                          ò
                                                                                                                                                ōλ
                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

Search completed: September 27, 2001, 14:22:24 Job time: 120 sec

```
4444444444466666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-672-564-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-672-564-1
114.5
1114.5
1114.5
1114.5
1114.5
1114.5
1113.1
1113.5
1111.5
1111.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
1110.5
110.5
110.5
110.5
110.5
110.5
110.5
110.5
110.5
110.
Sequence 1, Appli
Sequence 33, Appl
Sequence 4, Appli
Sequence 4, Appli
Sequence 19, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 19, Appl
Sequence 10, Appl
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 10, Appli
                                                                                                                                                                                                                                            (without alignments)
533.212 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                   US-09-189-415A-2
2800
1 MPIGNLGNNVNGNHLIPPAP......GETAVSSVNAAPTPGFVRFV 549
                                                                                                                                                                                                               September 27, 2001, 14:20:24; Search time 21.2 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence Sequence 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                         GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5268270-2

US-08-66-635A-19

US-08-61-697-10

US-08-194-468-2

US-08-302-832-2

US-08-302-832-2

US-08-301-98-2

US-08-517-697-2

US-08-728-470-2

US-08-728-470-2

US-08-719-641-2

US-08-719-641-2

US-08-119-641-3

US-08-186-495-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-185-432-19
US-08-194-290-7
US-08-614-377A-7
US-09-142-648B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197339 seqs, 20590346 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  122.5
118.5
11.18.1
11.7.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117.5
117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               138.5
138.5
134.5
134.5
132.5
132.5
122.5
                                                                                                                                                                                                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
```

```
Sequence 8, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16;
                                                                                                                                                                                                                                                                                                                                                           Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CERDEACH INFOUNDATION:

APPLICANT: IZUMI, YOSHIYA
APPLICANT: IZUMI, YOSHIYA
APPLICANT: IZUMI, YOSHIYA
APPLICANT: IZUMI, YOSHIYA
APPLICANT: IZUMI, WHISUMI
APPLICANT: IZUMI, WHISUMI
APPLICANT: IZUMI, WHOSHIM
APPLICANT: TIO, MAKOTO
ITILE OF INVENTION: GENE ENCODING ENDOGLYCOCERAMIDASE ACTIVATOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blich, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STREET: P.O. Box 747
CITY: Ealls Church
STREET: P.O. Box 747
CITY: TIGHINIA
COUNTRY: USA
ZIP: Z2040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: TBM PC compaTIBLE
COMPUTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                        6,000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 580;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.0%; Score 141; DB 2; Length 58
23.9%; Pred. No. 0.00043;
ive 45; Mismatches 126; Indels
                                                    US-09-323-433A-8

US-09-323-433A-8

US-08-728-470-10

US-08-708-70-70

US-08-918-35-2

US-08-928-361B-30

US-08-700-651-5

US-08-928-361B-6

US-08-928-361B-6

US-08-928-361B-6

US-08-928-361A-6

US-08-35-10661A-6

US-08-35-10661A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1, Application US/08672564 Patent No. 5824503 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MArc S.
REGISTRATION NUMBER: 32,181
REFERENCE/CDCKET NUMBER: 1422
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8050
TELEFAX: (703) 205-8050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELER: (703) 200
TELER: 248345
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 5.0
Best Local Similarity 23.9
Matches 84; Conservative
single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
```

g ò

g ò

ò

à

```
388 RRNQPAEQTTTTTTHTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSS 447
                                                                                                                                                                                                                                                      448 EVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQ 507
                                                                                    AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
    AESATK--DQLTQEAFKNPENQKVNID-----ANGNAIPSGELXDDIVEQIAQQAKE 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 2353;
                                                                                                                                                                                                       ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP STREET: Four Embarcadero Center, Suite 3400 CITY: San Francisco STATE: California COUNTRY: United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIAL.

COUNTRY: ULLIA.

ZIP: 94111-4187

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Ploppy disk

CORPUTER: IBM PC compatible

COPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.30

CURRENT APPLICATION NUMBER: US/08/913,942

TILING DATE: 2-DEC-1997

TILING DATE: 2-DEC-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: St. Geme, Joseph
APPLICANT: St. Geme, Joseph
APPLICANT: Barenkamp, Stephen J.
TILE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach Test Albritton & Herber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.9%; Score 138.5; DB 4;
22.0%; Pred. No. 0.0064;
tive 62; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: Vance, Dolly A.
REGISTRATION NUMBER: 39,054
REGISTRATION NUMBER: A-61053-1/RFT/RMS/DAV
RELEPHONE: (415) 781-1989
TELEPHONE: (415) 781-1989
TELEPK: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCIT/US96/4031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: PCT/US96/4031
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                        1570 TEVKIGAKTSVIKDHNGKLFTGKE 1593
                                                                                                                                                                                                                                                                                                                                               508 STYALLANSGGLRLGMGGLTSGGE 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 4, Application US/08913942 Patent No. 6200578 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2353 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 22.0
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 , MOLECULE TYPE: protein US-08-913-942-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
US-08-913-942-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                            1498
                 278
                                                                                                                                                                                                                                                                                                        qq
                                                                                                   δ
                                                                                                                                    g
                                                                                                                                                                                 δy
                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1163 LGGTANG------RNDTGTVINKDG-----LTITLANGAAAGTDASN 1198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1199 ------GNTISVTKDGISAĞNKEITNVKSALKTYKDTQNT------ADETQDKEFHA 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 AIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR--- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            179 ----TRILAKLDPDNHGGRQPKDVDTRSVGYGSASGIDDGVVSETHTST-TNSSVR--- 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----SDPKFWVSVGAIAAGLAGLAATGIAQALALIPEPDDPTTTDPDQAANA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPL----DILNTQIGPSAFRVEVQADGTHA 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.9%; Score 138.5; DB 4; Length 2353;
22.0%; Pred. No. 0.0064;
tive 62; Mismatches 207; Indels 171; Gaps
                                                                                                                       268 RQAVA-TVSYLDGPSAVTNGGEFTLNATVVPTPDSGQVQFTRDGEDVGAPVDLVN---GK 323
                                                                                                                                                                SAFRVEVQADGTHAAIGEKNGLEV-----SVTLSPQEWSSLQSIDTEGKNRFVFTGG 159
                                                                                                                                                                                                                                                  RGGSGHPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSET 219
                                                                                                                                                                                                                                                                            HTSTINSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALAL-----TPEPDDPTTTDP 271
                                                                                                                                                                                                                                                                                                                                                                           468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 LGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRD 65
                                                                                54 RNSMADSVDSRDIPGLPTNPSR--LAAATSETCLLGGFEVLHD----KGPLDILNTQIGP 107
                                  4 GNLGNNVNGNHLIPP-----APPLPSQTDGAARGGTGHLISSTGALGSRSLFSPL 53
                                                                                                                                                                                                                                                                                                                                                              | |: |: : : | | : | | 469 DPIASTARGS---PVTLTARLDPADARGTVQFKLGDVLLGGPVRVDANGVA 516
                                                                                                                                                                                            :: || :| | : || 324 ASLTQSLDTDGDXAYEAKFLGAEFFNPSSAAKTVTVTSQDIQTTTSVT-----
                                                                                                                                                                                                                                                                                                                                                                                                                    272 DQAANAAESATKDQLTQEAFKNPENQK--------VNIDANGNA 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: PEAK, Ian Richard Anselm APPLICANT: PEAK, Ian Richard Paul APPLICANT: JENNINGS, Michael Paul APPLICANT: JENNINGS, Michael Paul APPLICANT: WOXON, E. Richard TITLE OF INVENTION: NOVEL SURFACE ANTIGEN FILE REFERENCE: 065064/0128
CURRENT APPLICATION NUMBER: 1999-08-19
PRIOR APPLICATION NUMBER: PCT/AU98/01031
PRIOR APPLICATION NUMBER: GB 9726398.2
PRIOR SEQ ID NOS: 33
SOFTWARE: PATENTIN VET. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-377-155-33
; Sequence 33, Application US/09377155
; Patent No. 6197312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ) ORGANISM: Haemophilus influenzae US-09-377-155-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 2353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEO ID NO 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                           108
                                                                                                                                                                                                                                                                                                                                           220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤.
                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                               a
```

26;

Gaps

```
230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1497
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                 1163 LGGTANG--------RNDTGTVINKDG-----LTITLANGAAAGTDASN 1198
                                                                                            -----GNTISVTKDGISAGNKEITNVKSALKTYKDTQNT-----ADETQDKEFHA 1243
                                                                                                                                                                                                                                                                                       1344 VVKGSNGATATETDKKKVATVGDVAKAIND-AATFV-----KVENDDSATIDDSPTDDG 1396
                                                                                                                                                                                                                                                                                                                                                                                                          122 AIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR--- 178
                                                                                                                                                                                         ----TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTST-TNSSVR--- 229
                                                                                                                                                                                                                                                                                                                                                                                      AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                      388 RRNQPAEQTTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSS 447
                                                                                                                                                                                                                                                                                                                        AESATK--DQLTQEAFKNPENQKVNID-----ANGNAIPSGELXDDIVEQIAQQAKE 327
LGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRD
                                                               66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPL----DILNTQIGPSAFRVEVQADGTHA
                                                                                                                                                                                                                                                          ------SDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    448 EVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              & Herbert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------GG----NGITDNEKKRAASVKD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Barenkamp, Stephen I.
APPLICANT: St. Geme III, Joseph W.
TITLE OF INVENTION: Haemophilus Adhesion Proteins
WUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Flehr, Hobbach, Test, Albritton STREET: Four Embarcadero Center, Suite 3400 CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERNCE/DOCKET NUMBER: A-61053/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US/08/409,995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
CMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : : : : | : 1570 TEVKIGAKTSVIKDHNGKLFTGKE 1593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STYALLANSGGLRLGMGGLTSGGE 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 4, Application US/08409995
Patent No. 5646259
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA: APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X: USA
94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 24 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
US-08-409-995-4
                                                                                             1199
                                                                                                                                                                                           179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1498
                              g
                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                           ò
                                                                                                                                                        g
                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                             ò
                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1198 -----GNTISVTKDGISAGNKEITNVKSALKTYKDTQNT-----ADETQDKEFHA 1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1284 GLEKDIDGKIKLKVDNIDGNNLLIVDA-TKGASVAKGEFNAVITDATTAQGINANERGKV 1342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SVK 1513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1514 DVLNAGWNVRGVKPA-SANNQVENI-DFVAT---YDTVDFVSGDKDTTSVTVESKDNGKR 1568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------RNDTGTVINKDG-----LTITLANGAAAGTDASN 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 AIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR--- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        179 ----TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTST-TNSSVR--- 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPL----DILNTQIGPSAFRVEVQADGTHA 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----SDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          278 AESATK--DQLTQEAFKNPENQKVNID------ANGNAIPSGELXDDIVEQIAQQAKE 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   328 AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       448 EVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                        6 LGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRD 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           388 RRNQPAEQTITITITITYVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSS
                                                                                                                                                                                                                                                                                                                             Query Match
4.8%; Score 134.5; DB 1; Length 1912;
Best Local Similarity 22.0%; Pred. No. 0.01;
Matches 124; Conservative 62; Mismatches 207; Indels 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 Herbert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4, Application US/08685467
Patent No. 6060059
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Barenkamp, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rebr. Hobbach, Test, Albritton & Hell
STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----99-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508 STYALLANSGGLRLGMGGLTSGGE 531
INFORMATION FOR SEQ ID NO: 4.
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 amirr
                                                                                                                                                                                                                                                                                                                                                         Best Local Similary
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE: California COUNTRY: United States
                                                                                                                                                                                             double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                       ; TOPOLOGY: unknown
US-08-409-995-4
                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: dou
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1162 LGGTANG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
US-08-685-467-4
```

```
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         253 IAQALALTPE---
                                                                                                                                                                                                                                                                                                                                                                                  STATE: New York COUNTRY: U.S.A. ZIP: 10036-2711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                              New York
                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                      US-08-185-432-19
                                                                                                                                                                                                                                                                                                                                          STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GNTISVTKDGISAGNKEITNVKSALKTYKDTQNT-----ADETQDKEFHA 1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1243 AV--KNANEVEFV------GKNGATVSAKTDNNGKHTVTI--DVAEAKVGD 1283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1284 GLEKDTDGKIKLKVDNTDGNNLLTVDA-TKGASVAKGEFNAVTTDATTAQGTNANERGKV 1342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1162 IGGTANG------RNDTGTVINKDG------LTITLANGAAAGTDASN 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           388 RRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSS 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTST-TNSSVR--- 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------SDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANA 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    278 AESATK--DQLTQEAFKNPENQKVNID------ANGNAIPSGELXDDIVEQIAQQAKE 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          328 AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               448 EVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQ 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR--- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPL----DILNTQIGPSAFRVEVQADGTHA 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRD 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1912;
                                                          SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,467
FILING DATE: 22-JUL-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
ATTORNEX/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 3;
                                                                                                                                                                                                                                                               NAME: S11va, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOOKET NUMBER: A-61053-2/RFT/RMS
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.8%; Score 134.5; Di 22.0%; Pred. No. 0.01;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STYALLANSGGLRLGMGGLTSGGE 531
                                                                                                                                                                                                                                                                                                                                            TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 1912 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 22.0%
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SS: unknown
unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        , MOLECULE TYPE: protein US-08-685-467-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: un
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1497
```

```
28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2364 QELLNGQGLGMNGNGQRNGVGPGVLPGGLCGMGGLSGAGNGNSREQGLSPPYSNQSPPHS 2423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2177 GNGGNNGNGN----ASGKQSNQTAKQKAAKKLIEGSPDNGLDATGSLRK---ASSKK 2229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2230 TSAASKKAANLNGL--NPGQLT-----GGVSGVPGVPP---TNSAVQAAAAAAV 2275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          175 AEARTRILAKLDPDNHGGRQPKDVD----TRSVGVGSASGIDDGVVSETHTSTINSSVR 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56 SMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQ 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116 ADGTHAAIGEKNGLEVSVTL-SPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDI 174
                                                                                                                                                   APPLICAN: Xu, Tian APPLICANT: Xu, Tian APPLICANT: Xu, Tian APPLICANT: Xu, Tian APPLICANT: Matsuno, Kenji APPLICANT: Matsuno, Meriex PROTEINS, NUCLEIC ACIDS, AND TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PDDPTTTDPDQAANAAESATKDQLTQEAF----- 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 GNLGNNVNGNHLIPPAPPLPSQT--DGAAR-----GGTGHLISSTGALGSRSLFSPLRN 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4.8%; Score 133; DB 1; Length 2703; 19.8%; Pred. No. 0.024; tive 83; Mismatches 193; Indels 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATUR: SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        230 SD-----APKFWVSVGAIAAGLAGL----AATG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2276 AAMSHELEGSPVGVGMGGNLPSPYDTSSMYS----
                                                                                                                                                                                                                                                                                                                                       1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
Sequence 19, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
                                                                                                                APPLICANT: Busseau, Isabelle
APPLICANT: Diederich, Robert J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (212) 790-9090
(212) 869-8864/9741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18,872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 73
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: (212) 869-8864/97
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2703 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 19.8
Matches 119; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
```

```
431 QTAG---NAVNTTLTQADVTVTGNSSTTAVTVTQTA------AATAGATVAGRVN 476
                                                                                                                                                                                                                                                                                                                                           315 D--DIVEQIAQQAKEAGEVAR------QQAVESNAQAQQRYEDQHARRQEELQLSSGIG 365
                                                                                                                                                                                                                                                                                                                                                                                   477 GAVTITDSAAASATTAGKIATVTLGSFGAATIDSSALTTVNL------SGTG 522
                                                                                                                                                                                                                                                                                                                                                                                                                                  366 YGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTHTVVQQOTGGIPQHKVALMPQERR 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            523 TSL------GIGRGALTA----TPTANTLTLNVNGLT--TTGAITDSEAA----- 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         426 RFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDEVAADPG 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                482 YSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GLRLGMGG 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            135 LSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTRILAKLDPDNHGGRQ 194
                                                                                                                                                           195 PKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATGIA 254
                                                                                                                                                                                                                                                    255 QALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSGELX 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   561 -- ADDGFTTINIAGSTASSTIASLVAADATTLNISGDARVTITSH----TAAALTG
                                                                                                              341 LNTNTSGAAQTVTAGAGQNLTATTAAQAANNVAVDGGANVTVASTGVTS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Smit, John
APPLICANT: Smit, John
APPLICANT: Bingle, Wade H.
APPLICANT: Bingle, Wade H.
APPLICANT: No. 5976664ellini, John F.
TITLE OF INVENTION: EXPRESSION AND SECRETION OF
TITLE OF INVENTION: HETEROLOGOUS
TITLE OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.30 CUBRENT APPLICATION DATA:
FILING DATE: 12-MAR-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/194,290
FILING DATE: 09-FEB-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E: Fish & Richardson PC
225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 7, Application US/08614377A Patent No. 5976864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US 07, FILLING DATE: 09-JUNE-1992 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         526 LTSGG---ETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: Tsao, Y. Rocky
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : | | : | : | : | : | 670 SVNGGDGTDVLVANVNGS 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Massachusetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-614-377A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: E
                                                                                                                                                                                                       390
                                                                                                                                                        δ
                                                                                                                                                                                                Q
                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                               οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                            g
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24;
                                                                                                                                                                                                                                                         2590 SFHSGQMNPPSIQSSMSGSSPSTNMLSPSSQHNQQAFYQYLTPSSQHSGGHTPQHLVQTL 2649
2424 VQSSLALSPHAYLGSPSPAKSLPSLPTSPTHIQAMRHATQ------QKQFGGSNLNSLL 2476
                                                                                         2477 GGANGGGVVGGGGGGGGGGGGPQNSPVSLGIISPTGSDMGIMLAPPQSSKNSAIMQTIS 2536
                                                                                                                                                               2650 ------DSYPTPSPESPGHWSSSSPRSNSDWSECVQSPAANNLYISGGHQANKGSEAI 2701
                                                                                                                                   323 QQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGV 382
                                                                                                                                                                                                                          383 TTALHRRNQPAEQT------TTTTH----TVVQQQTGG-IPQHKVALM 420
                                                                                                                                                                                                                                                                                                                 421 PQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAE--VGGARNSL-----SAHQPEEHI 472
                                              ----KNPENQKVN---IDANGN-----AIPSGELXDDIVEQIA 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              230 TDNAAGVNLFTAYPSSGVSGSTL-----SLTTGTDT--LTGTANNDTFVAGEVAGAAT 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86 LGGFEVLHDKGPLDILN-----TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26 TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Smit, John
APPLICANT: Single, Wade H
APPLICANT: Bingle, Wade H
TITLE OF INVENTION: Bacterial surface protein expression
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFWHARE: Patentin PC-DOS/MS-DOS
SOFWHARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/10/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.7%; Score 132.5; DB 1;
21.0%; Pred. No. 0.0059;
tive 66; Mismatches 232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shlesinger, Arkwright & Garvey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 7, Application US/08194290 Patent No. 5500353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3000 South Eads Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC. TO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 59
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-5600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: Garvey, George A
REGISTRATION NUMBER: 17737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 1026 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 117; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           703-836-5288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-290-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Arlington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 09
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                2702 x 2702
                                                                                                                                                                                                                                                                                                                                                                                                     473 Y 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                            292
                                          ò
                                                                                  q
                                                                                                                                 δý
                                                                                                                                                                    g
                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

610

us-09-189-415a-2.rai

```
27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     425 TAVTVAQTAG---NAVNTTLTOADVTVTGNSSTTAVTVTQTA-------AATAGAT 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  309 PSGELXD--DIVEQIAQQAKEAGEVAR-----QQAVESNAQAQQRYEDQHARRQEELQ 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      360 LSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVAL 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                519 --SGTGTSL-----GIGRGALTA-----TPTANTLTLNVNGLT--TTGALTDSEAA- 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420 MPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 135 LSPQEWSSLQSIDT-EGKNRFVFTGGRGGS-----GHPMVTVASDIAEARTRILAKLDPD 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              249 AATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LIVGDILSGGAGTDVLNWVQAAAVTALPTGVTISGIETMNVTSGAAITLNTSSGVTGLTA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGL 248
                                                                                                                                                                                                                                                                                                                                                                                              230 İDNAAGVNLFTAXPSSGVSGSTL----SLTTGTDT--LTGTANNDTFVAGEVAGAAT 280
                                                                                                                                                                                                                                                                                                                                                                                                                                             86 LGGFEVLHDKGPLDILN------TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                    26 TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    476 VAADPGYSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                471 VAGRVNGAVTITDSAAASATTAGKIATVTLGSFGAATIDSSALTTVNL-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     THILLY DATE: A PROJUCT THOMAS F.; Halter, Roman; Pohlner, Johnnnes APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johnnnes TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM; NUMBER OF SEQUENCES: 6

CURRENT APPLICATION DATA:

FILING DATE: 01-JUL-1987
                                                                                                                                                                                                                                                                4.7%; Score 132.5; DB 4; Length 1026; 21.8%; Pred. No. 0.0059;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1507;
                                                                                                                                                                                                                                                                                                                 64; Mismatches 222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 128;
PRIOR FILING DATE: 1997-03-10
PRIOR APPLICATION NUMBER: US 07/614,377
PRIOR FILING DATE: 1996-03-12
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              664 TLGAGGSVNGGDGTDVLVANVNGS 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  520 RLGMGGLTSGG---ETAVSSVNAA 540
                                                                                                                                                                          ; ORGANISM: Caulobacter crescentus US-09-142-648B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.68;
                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 21.89
Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
5268270-2
;Patent No. 5268270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 1507
                                                                                                                                         1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO:2
                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5268270-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: Smit, John
APPLICANT: Smit, John
APPLICANT: Smit, John F.
APPLICANT: Bingle, Wade H.
TITLE OF INVENTION: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CAULC
FILE REFERENCE: 08106/002002
CURRENT APPLICATION NUMBER: US/09/142,648B
CURRENT FILING DATE: 1999-03-30
PRIOR APPLICATION NUMBER: PCT/CA97/00167
                                                                                                                                                                                                                                                                                                                                                                                            27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 309 PSGELXD--DIVEQIAQQAKEAGEVAR-----QQAVESNAQAQQRYEDQHARRQEELQ 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       360 LSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVAL 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 519 --SGTGTSL-----GIGRGALTA----TPTANTLTLNVNGLT--TTGAITDSEAA- 560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420 MPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    476 VAADPGYSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GL 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             249 AATGIAQALALIPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       425 TAVTVAQTAG----NAVNTTLTQADVTVTGNSSTTAVTVTQTA-------AATAGAT 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  || | :| : || || || 605 AAALTGITVTNSVGATLGAELGAGTTKAIVWGAGDDTVTVSSA 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       189 NHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGL 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                230 TDNAAGVNLFTAYPSSGVSGSTL----SLTTGTDT--LTGTANNDTFVAGEVAGAAT 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LIVGDILSGGAGIDVLNWVQAAAVIALPIGVIISGIETMNVTSGAAIILNISSGVIGLIA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                135 LSPQEWSSLQSIDT-EGKNRFVFTGGRGGS-----GHPMVTVASDIAEARTRILAKLDPD 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----GTTTVGANSAA---SGTVS---VSVANSST-----TTTGAIA--VTGG 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGGFEVLHDKGPLDILN-----TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT 134
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      26 TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               561 -----ADDGFTTINIAGSTASSTIASLVAADATTLNISGDARVTITSH-----T
                                                                                                                                                                                                                                                                                                                                                    Length 1026;
                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                  Query Match
4.7%; Score 132.5; DB 2;
Best Local Similarity 21.8%; Pred. No. 0.0059;
Matches 123; Conservative 64; Mismatches 222;
                                       08106/002001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 520 RLGMGGLTSGG---ETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                664 TLGAGGSVNGGDGTDVLVANVNGS 687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 7, Application US/09142648B Patent No. 6210948
              REGISTRATION NUMBER: 34053
REFREENCE/DOCKET NUMBER: 081
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
US-09-142-648B-7
                                                                                                                                                                                                                                                                                          US-08-614-377A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            390
```

g

ò

셤

g ò

g ò

ð

a ð 셤 ð g ò g ð q ò

ò

g

ò 음 õ g ò 쉽 ò qq ð g ò g ò g ö g ð g

```
103 TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGG 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           163 SGHPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTS 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               223 TINS----SVRSDPKFWVSVGAIA-AGLAGLAATGIAQALA-----LTPEPDDPTTTDP 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           272 DQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEV 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----TLTTLSSEPGQS 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                332 ARQQ-AVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRN 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               391 QPAEQTTTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGS---VASTHWSDSSS 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      428 -PSYPPITRSEYDYADHQNSG------SYYSHAAGQGSGLYSTFTYMNPAQR 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     234 PPTPPTTPKTDVQAGKVDLKREGRPLAEGGRQP-PIDFRDVDIGELS---SDVISNIETF 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------1SYSPFNLPHYR- 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    290 DVNEFDQYLPPNGHPGVPATHGQVTYTGSYGISSTAPTPATAGHVWMSKQQAPPPPPQQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 4.4%; Score 122.5; DB 4; Length 507; Best Local Similarity 19.1%; Pred. No. 0.015; Matches 75; Conservative 48; Mismatches 168; Indels 101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 507;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Protei
TITLE OF INVENTION: High Molecular Weight Surface Protei
TITLE OF INVENTION: of Wo. 5977336-Typeable Haemophilus
NUMBER OF SEQUENCES: 11
CONRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bidg. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                448 EVVNPYAEVGGARNSLSAHQP---EEHIYDEV 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : | : | : | : | : | 473 PMYTPIGDTSGVPSIPQTHSPQDWEQPVYTQV 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            350 РОАРОАРОАРРОООАРРООРОАРОООООЯН----
ATTORNEY/AGENT INFORMATION:
NAME: Digigilo, Frank 3.46
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10981
TELEPHONE: 516-742-4343
TELEPHONE: 516-742-4364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 10, Application US/08617697 Patent No. 5977336 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDUJUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DC
SOFTWARE: PatentIn Release **
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          391 ORTHIKTEQLSPSHYREQQQHSPQQ-
                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                        single
                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS:
TOPOLOGY: line
                                                                                                                                                                                                                                                                                                                                                                                            , MOLECULE TYPE:
US-08-860-635A-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-617-697-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                             22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         955 GRPVKPAPSPAANTASQAQKATQTDGAQIAKPQNIVV-----APPSPQANQ--AEELR 1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1056 QKAEAERKARELARQKAEEAS------HOANAKPKRRRAILPRPPAP 1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1099 VFSLDDYDAKDN-----SESSIGNLARVPIRMGRELINDYEEI------PL 1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                323 QQAKEAGEVARQQAVESNAQAQQRYED-----QHARRQEELQLSSGIGYGLSSALIVAG 376
                                                                                                                                                                                                                                                                                                        153 RFVFTGGRGGSG--HPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASG 210
                                                                                                                                                                                                                                                                                                                                        377 GIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRR------ 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         427 -FS----DRRDSQGSVASTHWSDSS-----SEVVNPYAEVGGARNSLSAHQPE 469
                                                                                                                                     757 TGYVTCNTGNLSDKAL-----NSFDXATKINGNVNLNONAALVLGKAALWGKIQ-- 805
                                                                                                                                                                                            93 HDKGPLDILNTQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKN 152
                                                                                                                                                                                                                                                   806 ------GQGNSRVSLNQHSKWHLTGDSQVHNLSLADSHIHLNNASDAQSANKY 852
                                                                                                                                                                                                                                                                                                                                                                                                                      211 ID--DGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAAT----GIAQALALTPEP 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       906 LDLFDA------SSVQDRSRLFVSLANHYVDLGALRYTIKTENGITRLYNPYAGN 954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 264 DDPTTTDPDQAANAAESATKDQLTQEA-FKNPENQKVNIDANGNAIPSGELXDDIVEQIA 322
                             Indels 142; Gaps
                                                                                    34 TGHLISSTGALGSRSLFSPLRNSMADSVD-SRDIPGLPTNPSRLAAATSETCLLGGFEVL 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Koopman, Peter
APPLICANT: Koopman, Peter
TITLE OF INVENTION: SOX-9 GENE AND PROTEIN AND
TITLE OF INVENTION: USE IN THE REGENERATION OF BONE OR CARTILAGE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
  Best Local Similarity 21.4%; Pred. No. 0.027;
Matches 104; Conservative 60; Mismatches 180;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E: Scully, Scott, Murphy & Presser
400 Garden City Plaza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIOM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,635A
FILING DATE: 29-MAY-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM9714
FILING DATE: 29-NOV-1994
APPLICATION NUMBER: AU PM9835
FILING DATE: 05-DEC-1994
APPLICATION NUMBER: PCT/AU95/00799
FILING DATE: 29-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 19, Application US/08860635A Patent No. 6143878 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 400 Garder
CITY: Garden City
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1139 EELEDE 1144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               470 EHIYDE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-860-635A-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
```

```
NAME: Relter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERNICE/DOCKET NUMBER: P41 9672
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)-546-4737
TELEPHONE: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
                                                               Sequence 2, Application US/08194468
Patent No. 5750336
                                                                                                     GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 19.0%
Matches 101; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein US-08-194-468-2
                                                                                                                                                                                                                                                                                CITY: Los Angeles
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 10 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KNG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 VDSR-----
                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                               US-08-194-468-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1278 --VESSSGSVTLVATGATLAVGNISGNTVTITADSGKLTSTVGSTINGTNSVTTSSQSG- 1334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----DIEGTI-----SGNTVNVTASTGDLTIGNSAKVEAKN 1366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1367 GAATLTAESGKLTTQTGSSITSSNGQTTLTAKDSSIAGNINAANVTLNTTGTLTTTGDSK 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1046 NISGFNKAEIT-AKNGSDLTIGNASGGNADAKKVTFDKVKDSKISTDGHNVTLNSEVKTS 1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1220 GNVNITASGNTLKVSNITGQDVTVTADAGALTTTAGSTISATTGNANITTKTGDINGK-- 1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303 ANGNAIPSGEL--XDDIVEQIAQQAKEAGEVARQ-----QAVESNAQAQQRYEDQHARRQ 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        412 IPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEH 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----SGPV-----TGRLIGTPGQG 505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : : | | : | | | : | : | : | | : | 30 ANINIKGVVKLGDINNKGGL--NITINASGTQKTIING--NITNEKGDLNIKNIKADAEI 985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          147 DTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTRIL----AKLDPDNHGGRQPKDVDTR 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       202 SVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA----IAAGLAGLAATGIAQAL 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            356 EELQLSSG-IGYGLSSALIVAGGIGAG---VTTALHRRNQPAEQTTTTTTTTTTVVQQQTGG 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            878 LTLGNISVEGNLSLTGANANIVGN-----LSIAEDSTFKGEASDNLNITGTFTNNGT 929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      50 FSPLRNSMADSVDSRDIPGLPINPSRLAAATSETCLLGGFEVLHDKGPLDILN-----T 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MPIGNL---GN-----NVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSL 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OIGPSAFRVE----VQADGTHAA--IGEKNGLE------VSVTLSPQEWSSLQSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     258 ALTPEPDDPTTTDPDQAANAAESATKDQLTQE-AFKNPENQKVNI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        506 IQSTYALLANSGGLRLG-----MGGLTSGGETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Pred. No. 0.2;
84; Mismatches 275;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           y Match 4.2%; Score 118.5;
Local Similarity 19.4%; Pred. No. 0.2
                                                                                                                                                  US PCT/US93/02166
                                                                                                                                                                 FILING DATE: 16 MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: BELKSTESSEC, JELTY W
RECISTRATION NUMBER: 22.651
REFERENCE/DOCKET NUMBER: 1038-557
TELECHONE: (703) 415-0810
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 amino acids
                                                   CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 05-0CT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/03
              APPLICATION NUMBER: US/08/617,697 FILING DATE: 01-APR-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       472 IYDEVAADPGYSVIQNFSG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 124; Conservative
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
TITLE OF INVENTION: COMPONDIVE GENES
TITLE OF INVENTION: COMPONDIVE GENES
TITLE OF INVENTION: RESPONSIVE GENES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----TSTINSSVRSDP----KFWVSVGAIAAGLAGLAATGIAQALALIPEPDDPTTTDPDQ 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAFPIDIKNISVITVPNMSQLQISVGIVP-----TQAIATGP-----TADPE- 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AANAAESATKDQLTQE-----AFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQ-QA 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----SQTGGQQM-----GATGVNPQLASKQSMVNSL 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---AATSET----CLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQA-----DGTHAAIGE 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----LEVSVTLSPQ--EWSSLQSIDTEGKNRFVFTGGRGGSGH 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      218 GRGAGMPYPAPAMQGATSSVLAETLTQVSPQMAGHAGLNTAQAGGMTKMGMTGTTSPFGQ 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      162 TSSPATSQTGPGICMNANFNQTHP----GLENSNSGHSLMNQAQQGQAQVMNGSLGAAGR 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 IPNGEL-SLLNSGNLVPDAASKHKQLSELLRGGSGSSINP--GIGNVSASSPVQQGLGGQ 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MPIGNLGNNVNGNHLIPPAAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     102 AQGQPNSTNMASLGAMGKSPLNQGDSSTPNLPKQAASTSGPTPPASQALNPQAQKQVGLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PMVIVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETH----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.2%; Score 118; DB 1; Length 2441;
19.0%; Pred. No. 0.43;
tive 70; Mismatches 177; Indels 184;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----DIPGL-----PTNPSRLA----
                                                                                                                NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark STRREET: 444 South Flower Street, Suite 2000
                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
```

Db 394 GDIALIDGNINAQGSGDIAKTGEFVETSGHDLFIKDNAIVDAKEMLLDFDNVSINAE 450 QY 190 HGGRQPKDVDTRSVGVG-SASGIDDGVVSETHTSTTNSSVRSDPKFW	OY 288 QEAFKNDENORVNIDDANGNAIPSGELXDDIVEQIAQAKEAGEVARQQAVESNAQAQRY 347	SUL - 08 Pat GE	CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/302,032 FILING DATE: 16-EEP-1994 CLASSIFICATION: 435 PRIOR APPLICATION DATA: APPLICATION NUMBER: GB 9205704.1 FILING DATE: 16-MAR-1992 PRIOR APPLICATION DATA: APPLICATION NUMBER: GB 9205704.1 FILING DATE: 16-MAR-1993 ATTORNEY/AGENT INFORMATION: NAME: BEFERICESSEN: Jeffy W REGISTRATION NUMBER: 22,651 REFERENCE/OCKET NUMBER: 22,651 REFERENCE/OCKET (703) 415-0810 TELECOMMUNICATION INFORMATION: TELECOMMUNICATION SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
Db 349	RESULT 14 US-08-038-682-2 Sequence 2, Application US/08038682 Sequence 2, Application US/08038682 Patent No. 5549897 GENERAL INFORMATION: GENERAL INFORMATION: TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS TITLE OF INVENTION: OF NON-TYPEABLE HAEWOPHILUS NUMBER OF SEQUENCES: 8 CORRESPONDENCE ADDRESS: ADDRESSEE: Shoemaker and Mattare, Ltd STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza STREET: Bldg.1 CITY: Arlington STATE: Virginia	COMPTRY: U.S.A. IP: 22202-0286 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM FC comparable ATORNEY AGENT INFORMATION: NAME: BERKSTRESSER, JERRY W REGISTRATION NUMBER: 22,651 REFERENCE/DOKET NUMBER: 1038-293 TELEPHONE: (703) 415-0810 TELEPHONE: (703) 415-0813 INFORMATION FOR SEO ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 1536 amino acids TYPE: amino acid STRANDEDNESS: single MOLECULE TYPE: protein US-08-038-682-2	Query Match 4.2%; Score 117.5; DB 1; Length 1536; Best Local Similarity 19.7%; Pred. No. 0.23; Pred. No. 0.23; Adaps 26; Matches 129; Conservative 73; Mismatches 239; Indels 213; Gaps 26; Oy 37 LISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLGGF

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     592 -----TITSGNOKGFRFNNVSLNGTGSGLOFTTKRTNKYAITNKFEGTLNISGK 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     348 EDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPA-----EQTTT---- 398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               641 VNISMVLPKNESGYDKFKGRIYWNLTSLNVSESGEFNLTIDSRG-----SDSAGTLTQ 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  694 PYNLNGISFNKDTTFNVERNARVNFDIKAPIGINKYSSLNYASFNGNISVSGGGSVDFTL 753
                                                                                                                                                                                                                                                                                                                                                                                                                                           274 NVRAATIRNQGKLSADSVSKDKSGNIVLSAKEGEAEIGGVISAQNQQAKGGKLMITGDKV 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141 SSLOSID----TEGKNRFVFTGG-RGGSGHPMVTVASDIAEARTRILAKLDPDN----- 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                394 GDIALIDGNINAQGSGDIAKTGGFVETSGHDLFIKDNAIVDAKEWL---LDFDNVSINAE 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         190 HGGRQPKDVDTRSVGVG-SASGIDDGVVSETHTSTTNSSVRSDPKFW------ 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      555 VDVHKN----ISLGAQGNI------591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      399 TTTHTVVQQQTGGIPQHK-----VALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVN 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       452 PYAEVGGARNSLSAHQPEEH--IYDEVAADPG---YSVIQ-----NFSGSGPVTGRL 498
                                                                                                                                                                                                                                                                                                                                    223 VISVNG--GSISLLAGQKITISDIIN------PTITYSIAAPENEAVNLGDIFAKGGNI 273
                                                                                                                                                                                                                                                                                                                                                                                                  90 ------EVLHDKGPLDILNTQIGPSAFRVEVQADGTHAAIGE------ 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          236 --vsygaiaaglaglaatgiaqalaltpepddptttdpdqaanaaesatkdqlt----- 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        288 QEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRY 347
                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                 37 LISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLLGGF----- 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   126 -----SVTLSPQE----WRGLEV-----SVTLSPQE----W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         451 TAGRSNTSEDDEYTGSGNSASTPKRNKEKTTLTNTTLESILKKGTFVNITANQRIYVNSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      499 IG-----TPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNAAPTPGPV 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 213;
                                                                                                                                                                                                    Length 1536;
                                                                                                                                                                                               Query Match
4.2%; Score 117.5; DB 1;
Best Local Similarity 19.7%; Pred. No. 0.23;
Matches 129; Conservative 73; Mismatches 239;
; LENGTH: 1536 amino acids
; TYPE: amino acid
STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-832-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

Search completed: September 27, 2001, 14:21:04 Job time: 40 sec

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

September 27, 2001, 14:20:24 ; Search time 25.62 Seconds
(without alignments)
1632.312 Million cell updates/sec Run on:

US-09-189-415A-2 2800 1 MPIGNLGNNVNGNHLIPPAP......GETAVSSVNAAPTPGPVRFV 549 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 Total number of hits satisfying chosen parameters:

219241 seqs, 76174552 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_68:* Database :

1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	probable transloca hemolysin A - Serr hypothetical prote hypothetical prote hypothetical prote collar protein iso puff 74E protein iso puff 74E protein iso puff 74E protein iso puff 74E protein prote submaxillary mucin hypothetical prote aggreean precursor iga-specific metal hypothetical prote related to C2H2 zi probable sensory h probable flagellar hypothetical prote proteoglycan core gene pipsqueak pro mucin, submaxillar hypothetical prote proteoglycan core gene pipsqueak pro mucin, submaxillar hemagglutinin/hemo hemagglutinin/hemo hemagglutinin-like probable zinc-fing paracrystalline su ascites sialoglyco flagellar hook-ass S-layer protein -
	E86045 A28182 C71513 C71513 C71513 C71513 T24583 T14583 T14583 T45233 T42317 T42317
DB	
Length	558 1608 1809 2830 2830 2233 1829 1829 1532 2132 1532 2132 1674 1674 1674 1995 1995 1995 1995 1995 1995 1995 199
ت نه	40000000000000000000000000000000000000
Sc	
Result No.	112 E 4 8 9 C 8 8 0 112 E 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

hypothetical prote probable RTX famil	hypothetical prote hypothetical prote	notch protein - fr hypothetical prote	A-alpha Z 4 protei	probable adhesin P	ALR protein homolo	neurogenic locus m	CREB-binding prote	hypothetical prote	aggrecan precursor	regulatory protein	high-molecular-wei	xanthomonapepsin (
085912 885547	F75477 H96597	A24420 T16509	7271	3135	2687	3106	3828	0365	9806	¥42020 -	143855	104900
B B	F7	AZ TI	D3	8 H 8	11	3 A3	11	T 0	L A3	2 A4	3 A4	S,
1528 5188	635 2	2703	930	1018	2422	1596	3190 2	1280	2415	618	1536	827
4.7	4.7	4.7	4.7	4.7	4.7	4.6	4.6	4.6	4.6	4.6	4.6	4.6
132	131.5 131.5	131 130.5	130.5	130.5	130.5	130	130	129	129	128.5	128.5	128
30	33 33	34 35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

RESULT 1 BE86045 Probable translocated intimin receptor protein tir [imported] - Escherichia coli (str probable translocated intimin receptor protein tir [imported] - Escherichia coli (str C;Species: Escherichia coli C;Species: I6-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001 C;Accession: E86045 R;Perna, N.T.; Plunkett III, G; Burland, V; Mau, B; Glasner, J.D.; Rose, D.J.; May Iller, L.; Grotbeck, E.J; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda Nature 409, 529-533, 2001 A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7. A;Recession: E86045 A;Recession: E86045 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-558 <2To> A;Residues: 1-558 <2To> A;Cross references: GB:AE005174; NID:g12518449; PIDN:AAG58825.1; GSPDB:GN00145; UMGP: C;Genetics: C;Genetics: A;Gene: tir	; Score 1528.5; DB 2; Length 558; ; Pred. No. 6e-88; 60; Mismatches 138; Indels 57; Gaps 11;	IGHLISSTGALGSRSLFSPLRNSMADS 60 - - - - - - - - - - - - -	CHDKGPLDILNTQIGPSAFRVEVQADG 118 - CHDHGPLDTLNRQIGSSVFRVETQEDG 116	NRFVFTGGRGGSGHPMVTVASDIAEAR 178 ::	DDGVVSETHTSTINSSVRS 230 : : : SELRESNSGAENTTETQTSTSTSSLRS 227	DDPTTTDPDQAANAAESATKDQLTQEA 290 : :	DAKEAGEVARQOAVESNAQAOORYEDQ 350 	FALHRRNQPAEGTTTTTHT 403 : AALHRRNQPVEQTTTTTTTTTSART 407
RESULT 1 BE86045 PERSOLLS PERSOLLS PERSOLLS PERSOLS PERSONS 54.6% Similarity 56.2% 7; Conservative	1 MPIGNLGNNVNGNHLIPPAAPPLFSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS	61 VDSRDIPGLPTNPSRLAMATSETCLLGGFEVLHDKGPLDILNTQIGPSMFRYEVQADG 	119 THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR 178 	179 TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRS	1 DPKFWVSVGAIAAGLAGLAATGIAQALALIPEPDDPTTTDPDQAANAAESATKDQLTQEA 	1 FKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQORYEDO	351 HARROBELOLSSGIGYGLSSALIVAGGIGAGYTTALHRRNQPAEQTTTTTHT 403 : ::	
RESULT 1 E86045 probable tr C; Species: C; Date: 16- C; Accession R; Perna, N., 111er, L.; Nature G9, A; Title: G9, A; Title: G9, A; A	Ouery Match Best Local Matches 32	Qy	Qy 63	Qy 115	Qy 175 Db 177	Oy 231 Db 228	Qy 291 Db 288	Qy 35. Db 348

RESULT 3 C71513 hypothetical protein CT456 - Chlamydia trachomatis (serotype D, strain UW3/Cx) C; Species: Chlamydia trachomatis C; Species: Chlamydia trachomatis C; Species: Chlamydia trachomatis C; Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999 C; Accession: C71513 R; Stephens, R.S; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche Science 282, 754-759, 1998 A; Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t A; Reference number: A71570; MUID:99000809 A; Reference number: A71570; MUID:9900809 A; Accession: C71513 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-1005 <arn> A; Respidues: 1-1005 <arn> A; Respidues: 1-1005 <arn> A; Repeinmental source: serotype D, strain UW-3/Cx C; Genetics: A; Gene: C7456</arn></arn></arn>	Ouery Match Duery Match Best Local Similarity 22.8%; Pred. No. 0.14; Matches 145; Conservative 65; Mismatches 234; Indels 192; Gaps 33; Qy 4 GNGNNVGNHIPPAPPL-PSQTDGAARGGTGHLISSTGALGSRSLFSPLR-54 Db 431 GNLPNTVIINNKFKTCVAXGPWNSQEASSGYTPSAWRRCHRVDFGGIFEKANDFNKINWG 490 Qy 55NSMADSVD-SRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGFS 108 Qy 55NSMADSVD-SRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGFS 108 Qy 55	AFRVEVQADGTHAAIGEKNGLEVSVT-LSPQEMSSLQSIDTEGKNRFVFTGGRGGSGHPM	OY 215 VVSETHTSTINSSVRSDPKFWVSV	OY 347 YEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGYTTALHRRNQPAEQTTTTT 4 01	Qy 503 GQGIQSTYALLANSGGLRGGGGTSGGETAVSSVN 538
QY 404 VVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSS-SEVVNPYAEVGGARNS 462	el, E.; Braun, V. characterization of the hemolysin determinant of Serratia and Sal S2 1988 A28182; MUD: 88257037 A CPOO> GB:M22618; NID: 9340726; PIDN: AAA50323.1; PID: 9556420 GB:M22618; Nover 150.5; DB 2; Length 1608; rity 20.9%; Pred. No. 0.25;	; Conservative 84 LISSTGALGSRSLESPLRNE :: :: TYDDAKGEGGTQRSNSSASQP :FEVLHDKGPLDILLNTQIGPE :: :: GDIRLDQASDKQSESRE	QY 121 AAIGEKNGLEVSVTLSPQEMSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEA-RT 179 b 1123 AQVGNISG-QQGVELKAGRDLTLQGTDVKSQGDVSLSAGN	293 293 1272 328 1332 388	437 VASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPQYSVIONESGSGPVTG

```
A; Accession: T34434
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Ross references: EMBL: U80846; PIDN: AAC70890.1; GSPDB: GN00028; CESP: K06A9.1a
A; Cross references: Expr. train Bristol N2; clone K06A9
C; Genetics: Strain Bristol N2; clone K06A9
A; Genetics: A; Map position: X
A; Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical protein T06D8.1 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C; Accession: T24583
B; Palmer, S.
Submitted to the EMBL Data Library, April 1995
A; Reference number: 219909
A; Accession: T24583
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-1829 < WILL>
A; Cross-references: EMBL: Z4930; PIDN: CAA88964.1; GSPDB: GN00020; CESP: T06D8.1
A; Experimental source: clone T06D8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    946 TVSDSTSSGSTVTVGSTEGSSSPIPSTSQNTNPSTSSGSSMSTQTPQSSQSTSPVESSTS 1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STE-----TPGSTGSTVTKP----STVSGSASSGSTATMGSTEASSTSGGSSTSPNPSQS 1099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1244 SGS-----STSGSSTPTPINP 1279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : |: | :: | :: | 1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1006 GATSSSGSPGTTLTSISPSPSPSSTI------GSSQGSTSPVV---STISQG 1048
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STGSTVTQPSTGSGSTSTSGEITSQGSTQTPRSSLSTSPAISTSTQQSVSTNSP---- 1213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GSTVTQPSTVRGSTSSGSTVTTGSTEGSST 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RLAAATS--ETCLLGGFEVLHDKGPLDILNT------QIGPSAFRVEVQAD 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  118 GTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEA 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           178 RTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVS 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238 VGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANA--AESATKDQLTQEAFKNPE 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NQKVNI----DANGNAIPSGELXDDIVEQIAQQA---KEAGEVARQQAVESNAQAQQRYE 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        349 DQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQ 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQP 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMG---G 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        899 PSPSQSTTIGSTQGSTSPGISTT------SEEMTSQGSTQTPGSTGSTVTQPS 945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PLPSQ--TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPG----LPTNPS 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match 5.3%; Score 149; DB 2; Length 2232; Best Local Similarity 19.9%; Pred. No. 0.5; Matches 112; Conservative 71; Mismatches 238; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STSGSSEKTSPSSISPVPTSSPI 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      526 LTSGG--ETAVSSVNAAPTPGPV 546
                                                A; Reference number: 221525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            T24583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
hypothetical protein 21495 [imported] - Escherichia coli (strain O157:H7)
C; Species: Escherichia coli
C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C; Accession: D8564
R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew 111er, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A; Reference number: A85480; MUID:21074935; PMID:11206551
A; Accession: D85644
A; Accession: D85644
A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary
A;Molecule type: DNA
A;Molecule type: DNA
A;Cross.references: GB:AE005174; NID:g12514354; PIDN:AAG55616.1; GSPDB:GN00145; UWGP:214
A;Experimental source: strain 0157:H7, substrain EDL933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: Caenorhabditis elegans
C; Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C; Accession: T3434
R; Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A; Description: The sequence of C. elegans cosmid K06A9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              118 TVQDIFAGG-----SRGADEVIDTLPDGKNAVTDIVGKG-----LKATGKAVSDGAKATD 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68 LPTMGKGFAQSVRGTGEM---ARGLGDAMIQSPVKTG-ARILNEFSRMGLP-----GVA 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TSETCLLGGFEVLHDKGPLDILNTQ-----IGPSAFRVEVQADGTHAAIGEKNGL 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E-VSVTLSPQEWSSLQSIDTEGKNRFVFTGGRG----GSGHP-MVT--VASDIAEARTR- 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRS 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         228 MLTAGLEKKYIAAGMQPERATALAAEAVDKKMPDLFQA----GLITHSTVSAQGQS--- 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          231 DPKFWVSVGAIAAGLAGLAA--TGIAQA-----LALTPEPDDPTTTDPDQAANAAESA 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  280 ------AMAAADAVLNADYSELAQSPKFQQTFLSIDADPQHAQLTD-RQKMDLAKER 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282 TKDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKE--AGEVARQQAVES 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          340 NAQAQQRYEDQHARRQEELQLSSGIGY----GLSSALIVAGGIGA-----GVTTALHRR 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 NQPAEQTITITHTVVQQQTGGIPQ-HKVALMPQER------RRFSDR-RDSQG 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  443 RQAAEETAMRDAETV--QQDDAAPQPESVDPVAQQRESMQGMNREQLLEQYADADMATEG 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                436 SVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVT 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              501 DASAAHRREAASOLL-----NELDEQTKRQAVMNELKAKPRSELLEEYRLSQKE 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 LPSQTDGAARG--GTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAA 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 2806;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.3%; Score 149.5; DB 2; Best Local Similarity 20.8%; Pred. No. 0.64; Matches 113; Conservative 88; Mismatches 218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    hypothetical protein K06A9.la - Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GR 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Experimental
C; Genetics:
A; Gene: 21495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     130 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
```

qq

οy

g ò g õ g à q

ö

g à g

à

ö g ò Dp ò g

Gaps

57

```
Duff 74E protein - fruit fly (Drosophila melanogaster)
NyAlternate names: ecdysone-induced protein E74B; ets-related protein E74B
C;ppecies: Drosophila melanogaster
C;ppecies: Drosophila melanogaster
C;paccission: S04722; B34692
C;Accession: S04722; B34692
R;Janknecht, R:; Taube, W:; Luedecke, H.J.; Pongs, O.
R;Janknecht, R:; Taube, W:; Luedecke, H.J.; Pongs, O.
R;Janknecht, R:; Taube, W: Lugles
A;Title: Characterization of a putative transcription factor gene expressed
A;Reference number: S04722; MUID:89315191
A;Molecule type: DNA
A;Recession: S04722
A;Molecule type: DNA
A;Residues: 1-883 - AJAN>
A;Residues: 1-883 - AJAN>
A;Residues: EMBL:X15087; NID:g7513; PIDN:CAA33195.1; PID:g7514
A;Note: Gln-867 was also found
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             457 GGARNSLSAHQPEEHIYDEVAADPGYSVIONFSGSGPVTGRLIGTPGQGIQSTYALLANS 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         284 SYASGNSAATNVKSKPEDETKSSDPSISESSGFKDTDVNAENEASAASVDDIEKLKALEE 343
                                                                                                                                                                                                                                                                                                                                           167 MVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDD------GVV 216
                                                                                                                                                                                                                                                                                                                                                                                     170 LSTAALNDSTPHPRNLGSV-TNNSAGR--SDDGEESLYLGRLFGEDEEEDYEGELVGGVA 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       227 NACEVEGLTTDEPFGSNCFANEVEIGDDEEESEIAEVLYKODVDLGFSLDQEAI---ING 283
                                                                                                                                                                        PGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAAIGEK 126
                                                                                                                                                                                                                                                            NGL-----REVETGGRGGSGHP 166
                                                                                                                                                                                                                                                                                                   110 RSIWEQNLADLYDYNDLSLQTSPYANLPLKDGQPQPSNSSHLDLSLAALLHGFTGGSGAP 169
                                                                                                                                                                                                                                                                                                                                                                                                                               ------DPKFWVSVGAIAAGLAG 247
                                                                                                                                                                                                      - GAPTSSGQTSGSA----LG--EIHIDTASLDPGNANHSPLHPTSELDFFLTPHALQDQ 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----VSDMSPYPHHY-----PGYSYQASPSNGAP-----GTPGQHGQY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DPVASTSQAAAAFNENQAQRIVSETGEDLLSGEGISSKQNRNEAKNKDNDPEKADGDSFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |: |: | | | 464 VSDFEELQNSVGSPLFDLEDARKELDEMLQSTVPSYHHPHPHPHPHPHPHASMH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TTTTHTVVQQQTGGIPQ-HKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  574 ------AGGFHHGHHQGRMPRLNRSVSMERLQD---FATYFSPIPS-----MV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---QAANAAESATKD-------Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------HRRNQPAEQTT
                                                                                        NNVNGNHLIPPAPPLPSQ-TDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDI
                                                Indels 319;
        Length 1296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LTQEAFKNPENQKVNIDANGNAIP-----SGE-------
5.2%; Score ... 0.44;
19.9%; Pred. No. 0.44;
....e. 68; Mismatches 213;
                                                                                                                              HITGNSSVQTAALQDVQSTSAAATGAT--MVVGTG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LAATGIAQALALTPEPDDPT-TTDPD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    517 GGLRLGMGGLTSGGETAVSSVNAAPTPGP 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GSGANATLOPPPPPP 664
                                                                                                                                                                                                                                                                                                                                                                                                                                        S--ETHTSTTNSSVRS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DIVEQIAQQAKEAGEVARQQAV----
                                                       Conservative
                                   Similarity
                                                           149;
                 Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 613 GG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                286
                                                           Matches
                                                                                                                                                                                     29
                                                                                                                                                                                                                                                                       127
                                                                                                                                                                                                                                                                                                                                                                                                                                            217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                             g
                                                                                                                                                                                       δ
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                    δŽ
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          function
                                              A;Gene: CESP:T06D8.1
A;Map position: 2
A;Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A,Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
A; Residues: 1-1296 <MCG>
A; Residues: 1-1296 <MCG>
A; Cross-references: EMBL:AF070064; NID:93859888; PID:93859889; PIDN:AAC72898.1
                                                                                                                                                                                         19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTELSKDDQVTEASGEETTTAAATEASEETTTSAVTEGSGEETTVVAVVESSGEEPASSS 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             485 IQNFS------GSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSV 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | : | : | : | : | : | 378 TITAAATEASEETTISAVIEGSGEDTIVVAVVESSGEQPASSSTSIPT-ELSKDDQVTEA 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGGIGAGVTTALHRRNQPAEQTTTT-----TTHTVVQQQTGGIPQHKVALMPQERR 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    497 TEASGEETITA--AATEASEETITSAVTEGSGEETIVVAVVESSGEEPASSSTSIPTELS 554
                                                                                                                                                                                                                                                                                                                                                                                                                                            267 SAVTEGSGE-----DITVVAVVELSGEQPASSSTSIPTELSKDDQVTEASGEETITA 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :: |: | :| |: | :| 319 AATEASEETTTSAVTEG-SGEETTVVAVVESSGEEPASSSTSIPTELSKDDQVTEASGEE 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        282 TKDQLTQEAFKNPENQKVN------IDANG------NAIPSGELX--DDIVEQ 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321 IAQQAKEAGEVARQQAVESNAQAQQRYEDQ-----HARRQEELQLSSGIGYGLSSALIV 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    496
                                                                                                                                                                                                                                                                         172 GEETTTSAVTEASSEA----TTTPAGTEASGEET----TTSAVTEGS 210
                                                                                                                                                                                                                                                                                                                   GLEVSV-----TLSPQEWSSLQSIDTE-GKNRFVFTGGRGGSGHPMVTVASDIAEARTRI 181
                                                                                                                                                                                                                                                                                                                                                             266
                                                                                                                                                                                                                                                                                                                                                                                                         227
                                                                                                                                                                                                                                 GLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAAIGEKN 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RFSDRRDSQGSVAST-HWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LAKLDPDNHGGRQPKDVDTRSVGVGSASG------IDDGVVSETHTSTTNSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           228 VRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDP-----DQAANAAESA
                                                                                                                                                 Length 1829;
                                                                                                                                             Query Match
5.2%; Score 146; DB 2; Length 18;
Best Local Similarity 21.7%; Pred. No. 0.58;
Matches 118; Conservative 63; Mismatches 256; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: cnc
A;Cross-references: FlyBase:FBgn0000338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Map position: 3
Keywords: leucine zipper
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAAPT 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : ||
674 TSIPT 678
```

375

á

g ò

g

182

음

g ò

à

211

g à a ŏ

à

89

126 555

g

g ò

ò

g

612

456

397

348 463 362

285

expressed in the 20

Genetics:

Gaps

144

294 247 343 414

530 585

```
A;Cross-references: EMBL:299168; PIDN:CAB16292.1; GSPDB:GN00066; SPDB:SPAC8C9.04
A;Experimental source: strain 972h-; cosmid c8C9
                  A;Status: preliminary; translated from GB/EMBL/DDBJ
A;MoLecule type: mRMA
A;Residues: 1.1589 «JIA»
A;Cross-references: EMBL:AF016589; NID:q3057086; PID:q3057087; PIDN:AAC39250.1
C;Genetics:
A;Gene: SM1
C;Superfamily: pig submaxillary mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hypothetical protein SPAC8C9.04 - fission yeast (Schizosaccharomyces pombe) c; Species: Schizosaccharomyces pombe c; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999 C; Accession: T39141 B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, September 1997 A.Reference number: 221748 A.Reference number: 221748 A.Reference number: 221748 A.Reference preliminary; translated from GB/EMBL/DDBJ A.Reference number: A.Reference SPACSCA.Rejules: 1-647 <0L1> A.Reference SPACSCA.Rejules: 1-647 <0L1> A.Residues: RMBL: Z99168; PIDN: CAB16292.1; GSPDB: GN00066; SPDB: SPAC8C
                                                                                                                                                                                                                                                                                                                                                                                                           66 GTNVP----VSGAPVTPGSSAGSSGAPGTGGPGSETASPLSGAAGTSATGSRTSIPP--- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56 SMADSVDSRDIPGLPTNPSRLAAATSETCL----LGGFEVLHDKGPLDILNTQIGPS--- 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --SGAPVTPEPPLISTGASAGPPASSESTVTLPGATGTDVLRSGTSLPVSGGAVTPASSP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LAATGIAQALALTPEPDD-PITIDPDQAANAAESAIKDQLTQEAFKNPENQKVNIDANGN 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              307 AIP-SGELXDDIVEQIAQQAKEAG------EVARQQAVESNAQAQQRYEDQHARR 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    392 TRPSSGETETTVIESRVSGSSDEGLGTIGSTAGLMRTTRISVVVSGTTGPSSGETGSA-- 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | : | : | | | : | : | : | : | : | 450 VSEFRISGSLGKG-SETTVSIPGL-ARMIRISFGGSRTIRQSSGETGTIVIESRISGSP- 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYD 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SEGLGRIGSTAGL----TRITII---SVVGSATTEPSSRETETTVTE 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIDTEGKNRFV------FTGGRGGSGHPMVTVASDIAEARTRILAKLDPDN-HG 191
                                                                                                                                                                                                                                                                                                                                                           GNNVNGNHLIPPAPPLPSQTDGAA----RGGTGHLISS-----TGALGSRSLFSPLRN 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----AFRVEVQA---DGTHAAIGEKNGLEVSVTLSPQEWSSLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSATRAPSSAVTRAPVTYDDVSGISHSSSGRSRTIVIGSPSSVSSAEQIAPSLSTDGLEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192 GRQPKDVDTRSV----GVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               295 TTKISDVDARTIRPSYGALGATGSSIGEIGTTSTS-----PEF-TETSSFSVGLR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          355 QEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTHTVVQQQTGGIPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGG----LTSGG
                                                                                                                                                                                                                                                                                                        87; Mismatches 247; Indels 156;
                                                                                                                                                                                                                                                     Length 1589;
                                                                                                                                                                                                                                                  5.1%; Score 144; DB 2; 19.5%; Pred. No. 0.63;
                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 119; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        531 ETAVSSVNA 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ETRTTVIES 594
                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
R;Burtis, K.C.; Thummel, C.S.; Jones, C.W.; Karim, F.D.; Hogness, D.S. Cell 61, 85-99, 1990
A;Tille: The Drosophila 74EF early puff contains E74, a complex ecdysone-inducible gene A;Reference number: A90912; MUID:90199900
A;Accession: B34692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - uou
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     submaxillary mucin 1 - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000
C;Accession: T42233
R;Jiang, W.; Woitach, J.T.; Keil, R.L.; Bhavanandan, V.P.
Biochem. J. 331, 193-199, 1998
A;Title: Bovine submaxillary mucin contains multiple domains and tandemly repeated, A;Reference number: Z22092; MUID:98180948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---KGPL----DILNTQIGPSAFRVEVQADG----THAAIGEKNGLEVSVTLSPQEWSSL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QSIDTEGKNRFVFTGGRGGSGHPMV-----TVASDIAEARTRILAKLDPDNHGGRQP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EEISNKS-------PPVQEDEEESESVASDCREFKVLYNHLRQQQHH--HSP 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                196 KDVD-TRSVGVGSASGIDD--GVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATG 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IAQALALTPE------PDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 224 ITAAPTLOPOOHQOPMSDIEDEETLEDVDDADADVEADAEDEELLEQYQNGYDSPLDLSL 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NG-----NAIP-----SGELXDDIV--EQIAQQAKEAGEVARQQAVE---- 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        284 GGATSAAASAAAAASAVSRRRGRTYSGTESDDSAQCERARMRLKPERKAERSAAYKKSLM 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----SNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTT 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSD 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        445 SSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPV----- 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  39 SSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHD---- 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SSSLSLSSSSSSSSSSTSSATPTPVAS --- PVTPTSPPPAAAAPAEASPPAGAELQEDGQQA 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           495 --TGRLIGTPG-QGIQSTYALLANSGG-----LRLGMGGLTSGGETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | | :| | : | | : | | | ATTGSFPPSPADSGVSDVDS--SSSGGQPCADELKARLGMPPATSASAAAAAAAA 521
                                                                                                                                                                                                                                                                                                                                                              transcription regulation
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Score 144; DB 2; Length 883;
; Pred. No. 0.28;
68; Mismatches 197; Indels 208;
                                                                                                                                                 A;Wolecule type: DNA
A;Residues: 1-866,'Q', 868-883 <BUR>
A;Cross-references: GB:M37083; NID:9157309; PID:9157310
                                                                                                                                                                                                                                             A;Gener: FlyBase:Eip74EF
A;Cross-references: FlyBase:FBgn0000567
A;Map position: 3 74EF
C;Superfamily: ets DNA-binding domain homology
C;Keywords: alternative splicing; DNA binding; transcri
F;789-869/Domain: ets DNA-binding domain homology <FIS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  391 LLHIK---TEONTLLTPLOLOQQQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 h 5.1%;
Similarity 20.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 20.8%
Matches 124; Conservative
                                                                                                                                                                                                    A;Cross-refe
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

```
A; Molecule type: mRNA
A; Residues: 350-481, 'R', 483-506 <GLU1>
A; Cross-references: EMBL:X80279; NID:9673432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule RNAA
A;Molecule RNAA
A;Mosicule RNAA
A;Residues: 1-2132 <WAL>
A;Residues: 1-2132 <WAL>
A;Residues: 1-2132 <WAL>
A;Cross-references: GB:LD7049; NID:9678541; PIDN:AAC37670.1; PID:9191772
B;Granabo, H.; Gao, L.; Sugiyama, S.; Doege, K.; Kimata, K.; Yamada, Y.
Biochem. J. 300, 433-440, 1995
A;Atile: Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene structum
A;Reference number: S55329; MUID:95289972
A;Accession: S55329
A;Accession: S5532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       aggreean precursor - mouse (Species: Mus musculus (house mouse) (Spaces)                                                                                                                                                                                                                                                                                                                       20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            437 VGG----TTAEHPVQEAVTAPTETAHDFSKETTTASKRVSKHDKASAEKHKVARKP--- 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSV 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 375 AGGIGAGVTTALHRRNQ----PAE-----QTTTTTTTTTVVQQQTGGIPQHKVALMPQER 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 IAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTN----- 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : | | : : | | : : | 1 | : : | 1 | : 1 | : : | 1 | : : | 1 | : : | 1 | : : | 1 | : : | 1 | : : | 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 294
                                                                                                                                                                                                                                                                                                                                                                                                               114 VQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASD 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOEPGFTQTIVEKDADQVDEPLEPIASSALGTVEPPTDNK------PSASTSTA 141
                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      485 IQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRGMGGLTSGGETAVSSVNAA---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               199 ASITSKPATTSAAQPSSKVEENWAKATSQPITTAEKEIPELKPIEPEAIMISKEINTTHD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----ANGNAIPSG--ELXDDIVEQIAQQAKEAGEV
                                                                                                                                                                                                                                                 Query Match 5.1%; Score 141.5; DB 2; Length 047; Best Local Similarity 19.0%; Pred. No. 0.25; Aatches 103; Conservative 64; Mismatches 207; Indels 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSVRSDPKFWVSV---GAIAAGLAGLAATGIAQALALTPE--PDDP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   332 ARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSAL----
                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                  5.1%; Score 141.5; Di
19.0%; Pred. No. 0.25;
C;Genetics:
A;Gene: SPDB:SPAC8C9.04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                295 ENQKVNID-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   577 TPG 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      542 TPG 544
                                                                                                            A; Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Holecule type: DNA
A;Residues: 211-326 <WAR2>
A;Cossion: 211-326 <WAR2>
A;Cossion: 158123
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 211-240, WCTASLREWRYRSFWRHPQRNSPSRRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRQPTS', 'AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-240, WCTASLREWRHPQRNSPSRQPTS', AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-2410, WCTASLREWRHPQRNSPSRQPTS', AGGWGHAWPPQASSTWPGRAVWTCAALAGW', A;Cossion: 211-2410, WCTASLREWRHPQRNSPSRQPTS', AGGWGHAWPQASSTWPG A) introns: 253/1 C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; C;Keywords: cartilage; extracellular matrix F;1-19/Domain: signal sequence #status predicted <SIG> A;Cross-references: GB:U22901; NID:9886014
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues
A;Reference number: S50206; MUID:95035091 R;Watenabe, H.; Kimata, K.; Line, S.; Strong, D.; Gao, L.Y.; Kozak, C.A.; Yamada, Mature Genet. 7, 154-157, 1994
A;Title: Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the A;Reference number: 158123; MUID:95004579
A;Accession: 178532 29; GLDTSANGGYVSGI-----PSGGDGTETSANGGVEDVSGLPNGGEGLETSANGGVEDLGP 1149 965 VEDLSGLPSGEEGSETSTS-----GIE-----DISVLPTG-GES---LETSASG---- 1004 180 RILAKLDPDNHGGROPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVG 239 ATKDQLTQEAFKNPENQKVNIDANGNAIPSGE-----LXDDIVEQIAQQAKEAGEV 331 -RDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAA 122 4 GNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDS 63 Gaps 906 GDLSGLPSGGEITETSTSGAEETSGLPSGGDG-LETSTSGVDDVSGIPTGREGLETSASG ----DQAANAAES---IGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVAS---DIAEART Ouery Match 5.0%; Score 140; DB 1; Length 2132; Best Local Similarity 20.5%; Pred. No. 1.7; Matches 129; Conservative 81; Mismatches 241; Indels 178; F;44-135/Domain: immunoglobulin homology <IMM>
F;170-247/Domain: link protein repeat homology <IMM>
F;268-349/Domain: link protein repeat homology <INK1>
F;568-349/Domain: link protein repeat homology <INK2>
F;560-683/Domain: link protein repeat homology <INK3>
F;502-683/Domain: link protein repeat homology <INK4>
F;1922-2042/Domain: C-type lectin homology <ICH>
F;2049-2105/Domain: complement factor H repeat homology <FHD> 240 AIAAGLAGLAATGIAQALALTPEPDDPTTTDP----Ridiunoff, V.

Submitted to the EMBL Data Library, July 1994
A; Reference number: S51355
A; Status: preliminary
A; Molecule type: mRNA
A; CFVMSQRERPWAA' <GLU2>
A; Cross-references: EMBL; X00779

Db 878 TIKINHLSGNGHFHYLTDLAKNLGDKVLVKESASGHYQLHYQNKT-GEPNQEGL 930 212 DDGVVSETHTSTTNSSVRSDEKFWVSVGAIAAGLAGLAATGIAQALALTPEPD 264 1	PAO1) 2000 Hickey, M.J	adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa A;Title: Complete A82950; MUID:20437337 A; Accession: A83412 A; Conetics: A; Conetics: A; Conetics: A; Gene: PA1874	Query Match 4.9%; Score 137; DB 2; Length 2468; Best Local Similarity 21.5%; Pred. No. 3.2; Matches 135; Conservative 57; Mismatches 275; Indels 160; Gaps 28; Qy 2 PIGNIGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLR 54	103 TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRF 103 TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRF 1
Qy 332 ARQQAVESNAQAQORYEDDHARRQEELQLSSG	RESULT 12 A26039 IgA-specific metalloendopeptidase (EC 3.4.24.13) precursor - Neisseria gonorrhoeae (stra Niternate names: IgA protease; immunoglobulin Al proteinase C;Species: Neisseria gonorrhoeae; immunoglobulin Al proteinase A;Variety: strain MS11 C;Species: O5-Cct-1898 #sequence_revision 05-Oct-1988 #text_change 08-Dec-2000 C;Accession: A26039; S09386 R;Pohlner, J.; Halter, R.; Beyreuther, K.; Meyer, T.F. Nature 325, 458-462, 1987 A;Title: Gene structure and extracellular secretion of Neisseria gonorrhoeae IgA proteas A;Reference number: A36039; MUID: 87115823	A; Accession: A26039 A; Molecule type: DAOH A; Mote: the authors translated the codon AAG for residue 668 as Asn B; Haller, R.; Pohlner, J.; Meyer, T.F. EMBO J. 8, 2737-2744, 1989 A; Title: Mosalc-like organization of IgA protease genes in Neisseria gonorrhoeae generat A; Recence number: S09386; MuID: 90060036 A; Status: not compared with conceptual translation A; Molecule type: DNA A; Experimental source: strain MS11 A; Experimental source: strain MS11	C;Genetics: C;Genetics: C;Genetics: C;Superfamily: IgA-specific metalloendopeptidase C;Superfamily: IgA-specific metalloendopeptidase C;Superfamily: IgA-specific metalloendopeptidase C;Keywords: hydrolase; metalloendopeptidase; transmembrane protein C;Superfamily: IgA-specific metalloendeinge; transmembrane protein F;127-100main: signal sequence #status predicted F;188-1532/Product: immunoglobulin Al proteinase #status predicted F;1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted F;1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted Ouerv Match 4.9%: Score 137; DB 2; Length 1532;	Similarity 21.0%; Fred. No. 1.6; 2; Conservative 60; Mismatches LISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTN :

Db 498 GASGGTSYRYTHTHTHSHSHSGLGDVPELDSSGMFVELPATTTTTASASASLPPRNSFG 557 Oy 263	Status: preliminary; translated fro Molecule type: DNA Residues: 1-1829 <har> Cross-references: EMBL:AL031031; Pl Experimental source: strain A3(2) Genetics: Gene: SCOEDB:SC7C7.03 Query Match 4.9%; Sco</har>	vative 90; Mismatches 274; Indels 103; Ga IPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVD :	Qy 122 AIGEKNGLEVSVTLSPQEWSSLQSIDPEGKNERVFTGGGGGGHPMYASDIAEATRI 181
QY 251 TGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDAN 304 DD 740 AGSSYTVDSVAPATPVINPSNGTTLSGTAEPGSSVTLT-DGNGNPIGQ-VTADGS 792 QY 305 GNAIPSGELXDDIVEQIAQAKEAGEVARQQAVESNAQAQRYEDQHARRQEELQLSS 362 QY 305 GNAIPSGELXDDIVEQIAQAKEAGEVARQQAVESNAQAQRYEDQHARRQEELQLSS 362 I	RESULT 14 T51024 related to C2H2 zinc finger transcription factor D-Spl [imported] - Neurospora crassa NiAlternate names: protein B7F21.50 C; Species: Neurospora crassa C; Species: Neurospora crassa C; Species: Neurospora crassa C; Species: 21-0u1-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000 C; Accession: T51024 R; Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000 A; Reference number: 225286	A,Accession: T51024 A,Status: preliminary A,Nolecule type: DNA A,Nolecule type: DNA A,Residues: 1-770 <sch> A,Cross =references: EMBL:AL389901; GSPDB:GN00116; NCSP:B7F21.50 A,Experimental source: BAC clone B7F21; strain OR74A C,Genetics: A,App position: 6 A,Hap position: 6 A;Introns: 117/1</sch>	Ouery Match Best Local Similarity 19.1%; Score 136.5; DB 2; Length 770; Best Local Similarity 19.1%; Pred. No. 0.67; DP 1. Lindels 193; Gaps 20; Matches 99; Conservative 56; Mismatches 170; Indels 193; Gaps 20; QY 16 IPPAPPLESCT-DGARGGTGHLISSTGALGSRSLFSPLRNSMADSVDS 63 : : : :

Search completed: September 27, 2001, 14:21:42 Job time: 78 sec

PAGE BLANK (USPTO)

```
September 27, 2001, 14:20:24 ; Search time 14.89 Seconds
(without alignments)
1263.013 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P48434 gallus gall
P1532 serratia ma
P11536 drosophila
G61282 mus musculu
P09790 netsseria g
P52143 escherichia
P7897 rattus norv
P12021 sus scrofa
P3582 caulobacteri
C24762 drosophila
P46590 candida alb
P61727 drosophila
P61739 schizophyll
P6170 drosophila
P16112 homo sapien
G60106 xanthomonas
P16112 homo sapien
Q60299 homo sapien
P189769 drosophila
P189769 drosophila
P20105 drosophila
P21056 drosophila
P21056 drosophila
P21055 bordetella
P2106 drosophila
P2106 drosophila
P21055 soccharomyc
P26639 mus musculu
P12255 soccharomyc
P26639 mus musculu
P12858 saccharomyc
P26639 mus musculu
P12858 saccharomyc
P201286 mus musculu
P14445 saccharomyc
P39468 mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                   US-09-189-415A-2
2800
1 MPIGNLGNNVNGNHLIPPAP......GETAVSSVNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93435 seqs, 34255486 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SYN_DROME
MSB2_XEAST
TALI_MOUSE
YR83_MYCTU
HTF4_MOUSE
XIQ9_XEAST
CDP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HLYA_SERMA
BCA_MOUSE
IGA_NEIGO
YEPJA_ECOLI
PGCA_RT
APMU_PIG
SIAP_CAUCR
CSG_HALVO
ZEST_DROVI
MAZ4_SCHCO
MAZ4_SC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHP_DROME
E74A_DROME
MAPB_MOUSE
FHAB_BORPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CYAA_USTMA
Y552_HUMAN
CBP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGCA_HUMAN
XANP_XANS2
SOX9_PIG
SP4_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOX9_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22132
21532
21532
21050
20120
20120
20120
20120
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
20130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123.5
123.5
123.5
123.5
122.5
122.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33100843310088492
33100843310884331088492
                                                                                                                                                                                                                                                                                                         Title:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
9
```

2541 TALL_HUMAN Q99490 homo sapien 2175 1 HMCU_DROME P10180 drosophila 1411 TCOF_HUMAN Q13428 homo sapien 2421 TEGU_HSVEB P45481 mus musculu 538 1 DAC_ACTSP P39045 equine herp 538 1 DAC_ACTSP P39045 actinomadur 933 1 Y144-HUMAN P14157 homo sapien 1073 1 R183 DROME P9869 drosophila 1099 1 PLC1_CANAL 013433 candida alb 3828 1 TRX_DROVI Q24742 drosophila 2364 1 PGCA_BOVIN P1360B bos taurus 1140 1 W966_YEAST Q04893 saccharomyc	ALIGNMENTS	DARD; PRT; 494 AA.	<pre>13, Created) 5, Last sequence update) 8, Last annotation update) NR SOX-9.</pre>	SOX9. Galus gallus (Chicken). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		<pre>x M., Cartwright E.J., Pearl G., Scotting P.J., 1994) to the EMBL/GenBank/DDBJ databases. LOCATION: NUCLEAR (POTENTIAL). CONTAINS 1 HMG BOX.</pre>	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).) 563.1; 1) (ation; DNA-binding; Nuclear protein. 174 BMC BOX. 174 BOLY SPER. 1850 GLN/PRO-RICH. 401 POLY-GLN. 54942 MW; AF4ID483B5016349 CRC64;	5.6%; Score 157.5; DB 1; Length 494; 18.9%; Pred. No. 0.039; rative 48; Mismatches 142; Indels 127; Gaps	TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEMSSLQSIDTEGKNRFVFTGGRGG 162 	SGHPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTS 222	-
44.2 21745 4.2 21745 4.2 21745 4.2 24411 4.2 3421 4.2 1073 4.2 1073 4.2 2364 4.1 1160		STANDARD;	Rel. 33, Rel. 35, Rel. 38, FACTOR S	(Chicken) tazoa; Cho Aves; Neogr	. A	Rex M., Ca) C-1994) to AR LOCATION Y: CONTAINS	Constructions of the construction of the const	EMBL; U1253; AAB09663.1; HSSP; O05066; 1HRZ. InterPro; IPR000910; Franscription regulation; DNA_BIND 106 174 DOMAIN 215 220 DOMAIN 342 360 DOMAIN 394 401 SEQUENCE 494 AA; 54942	Similarity 18.	RVEVQADGTH2 : FKALQAD	'ASDIAEARTR' 	
120 119 118 118 117.5 117.5 117.5 117.5 117.5		T 1 CHICK SOX9_CHICK	10 - 01	K9. llus gallus karyota; Me chosauria;	ID=9C	<pre>Uwanogho D., Rex Sharpe P.T.; Submitted (DEC-1: -!- SUBCELLULAR: -!- SIMILARITY:</pre>	This SWISS-PR Detween the the European use by non- modified and entities requ or send an em	EMBL; U12533; AAB09663. HSSP; Q05066; 1HRZ. InterPro; IPR000910; - Pfam; PF00505; HMG_box; Transcription regulation DNA_BIND 106 174 DOMAIN 215 220 DOMAIN 342 360 DOMAIN 394 401 SEQUENCE 494 AA; 549	atch sal Simi 74;	103 TQIGPSAFRVEVQAD : 197 THISPNAIFKALQAD	163 SGHPMVTV 1 235 PPTPPTTP	
335 337 441 443 5		RESULT SOX9_CHI	01- 01- 15-	Gall Gall Euka:	SEC SEC	She Suk - 1 -	This bety the the mod ent.	EMI HS: Int Pfc Tre DON DON SEC	Query Ma Best Loc Matches	15	75	č

~

us-09-189-415a-2.rsp

```
85 LLGGFEVLHDKGPLDILNTQIGPSAFRVEVQA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E74B_DROME
                                                                                                                         121
                                                                                                                                                                              1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      328
                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                           d
                                                                                                                         ò
                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 292 DVNEFDQYLPPNGHPGVPATHGQVTTYSGTYGISSSASSPAGAGHAWMAKQQPQPPQPPA 351
                                                                 269 TDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEA 328
                                                                                                                         -----HTLPS-----TEREQGPAQQR 372
                                                                                                                                                                                                                                                                                                                                                               420 -----SYYSHAAGQSGSLYSTFT 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 141; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Poole K., Schiebel E., Braun V.; "Molecular characterization of the hemolysin determinant of Serratia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               33 GTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSR-----LAAATSETC 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA REQUIRES SHLB FUNCTION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J. Bacteriol. 170:3177-3188(1988).
-1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
                                                                                                                                                                                    329 GEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIG-----YGLSSALIVAGGIGAGV
                                                                                                                                                                                                                                               373 PHIKTEQLSPSHNSEQQQHPEQQQQQQQL----GYGSFNLQHYGFSYPPI-----
                                                                                                                                                                                                                                                                                                       383 TTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVAS--T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1608;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31 1608 HEMOLYSIN.
1608 AA; 165078 MW; D669B476FE7DAD51 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.4%; Score 150.5; DB 1; 20.9%; Pred. No. 0.42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84; Mismatches 236;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: OUTER MEMBRANE. SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
01-NOV-1990 (Rel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1608 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A., AND SEQUENCE OF 31-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hemolysis; Toxin; Outer membrane; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                           441 HWSDSSSEVVNPYAEVGGARNSLSAHQPEEH 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      452 YMNPTQRPMYTPIADTSGVPTIPQTHSPQQH 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-88257037; PubMed-3290200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M22618; AAA50323.1; -. PIR; A28182; A28182.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                               352 OPPAO-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Serratia marcescens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HEMOLYSIN PRECURSOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          marcescens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HLYA_SERMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOUTH THE TRANSCOORD OF THE TR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                 à
                                                                                                                         a
                                                                                                                                                                                    à
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
AQVGNISG-QQGVELKAGRDLTLQGTDVKSQGDVSLSAGN-----KVALQAAESTQT 1173
                                                                                                                                                                                                          1174 RKESKLSGNIDLGAGSSDSKEKTGGNLSAGGAFDIAKVNESATERQGATIASDGKVTLSA 1233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1332 AGDVTLNSGKDTRL-AGARVD-----ADSVQGKVGGDLHVESRKDVENGVKVDVDAGLS 1384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1385 HSNDPGSSITSKLSKVGTPRYAGKVKEKLEAGVNKVADATTDKYNSVARRLDPQQDTTGA 1444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1445 VS---FSKAEGKVILPATPAG-----EKPQGPLWDRGARTVGGAVKDSITGPAGRQG 1493
1066 VNAGGDIRLDQASDKQSESR---SGFNVKASAKGGFTADSKNFGAGFGGGTHNGESSSST 1122
                                                                                                                                                                                                                                                                                                                                                                                                                                 1272 HKDNWSLGIKANAKGGQTFNKDAGGKVDPNTGKDTHTLGAGLKVGVEQQDKTTHANTGIT 1331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EIPTQUE OR E74B OR GG285.

BIP74EF OR E74B OR GG285.

Drosophila melanogaster (Fruit fly).

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Mooptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    437 VASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTG 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               388 RRNQPAEQITITITHTVVQQQTGGIPQ-----HKVALMPQER-----RRFSDRRDSQGS 436
                                                                                                                                                                                                                                                                      ---GAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAA---ESATKDQLTQEAFK 292
                                                                                                                                                                                                                                                                                                                                                                              293 NPEN------QKVNIDANGNAIP----SGELXDDIVEQIAQQAK-----E 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-BERKELEY:
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adams M.D., Celniker S.E., I. P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          expressed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Burtis K.C., Thummel C.S., Jones C.W., Karim F.D., Hogness D.S.; "The Drosophila 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-related proteins."; Cell 61:85-99(1990).
                                                     AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGSGHPMVTVASDIAEA-RT
                                                                                                                                                             RILAKLDPDNHGGRQPKDVDTRSVGVGSASG-IDDGVVSETHTSTTNSSVRSDPKFWVSV
                                                                                                                                                                                                                                                                                                                           1234 NGKGDDALHLOGAKVSGGSAAL-------EAKNGGILLESAKNEQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E74E_DROME STANDARD; PRT; 883 AA.
P11536; Q9VVI7;
01-0CT-1989 (Rel. 12, Created)
01-0CT-1989 (Rel. 12, Last sequence update)
01-0CT-2000 (Rel. 40, Last annotation update)
ECDYSONE-INDUCED PROTEIN 74EFB (ETS-RELATED PROTEIN E74B).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A. MEDLINE=89315191; PubMed=2501755; Janknescht R., Taube W., Luedecke H.-J., Pongs O.; Janknescht R., Taube W., Luedecke H.-J., Pongs O.; "Characterization of a putative transcription factor ger in the 20-OH-ecdysone inducible puff 74EF in Drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | : | | : | | : | | 111 | 1494 | HLKVNADVVNNNAVGEQSA---IAGKNGVALQVGGQTQLTGGE 1533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  497 RL-----IGTPGQGIQSTYALLANSGGLRLGMGGLT--SGGE 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nucleic Acids Res. 17:4455-4464(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=90199900; PubMed=2107982;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mes by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Sutton G.G., Wortnan J.R., Yandell M.D., Zhang Q., Chen L.X.,
As Brandon R.C., Rogers Y.H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
As Brandon R.C., Rogers Y.H.C., Blazel R.G., Nelson C.R., Miklos G.L.G.,
Abril J.E., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
As Ballew R.M., Basua A., Baxendale J., Barrataroqiu L., Besaley E.M.,
Bescon K.Y., Benos P.W., Berman B.P., Brokaten P., Bortlar P.,
As Dorkow D., Botchan M.R., Bouck J., Brokaten P., Bortlar P.,
Burtis K.C., Busam D.A., Bulch H., Cadlau E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davise P.,
Burtis K.C., Busam D.A., Bulch H., Cadlau E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davise R.,
As Durbin K.J., Exangelista C.C., Ferraz C., Ferriera S., Felschmann W.,
Robiel C., Gabrielian A.E., Gary N.S., Galbart W.M., Glasser K.,
As Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
As Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
As Gladek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Mattel B., Kodire C.D., Karif C., Kravitz S., Kulp D., Lai Z.,
As Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
As Hasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
As Harris M.W., Murphy L., Muzphy L., Muzny D.M., Nelson D.L.,
As Dilali M., Pittman G.S., Pan S., Pollard J., Puri V., Resee M.G.,
As Blazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Sahen H.,
As Dienert K., Remington K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
As Dienert K., Sannders R.D.C., Scheeler F., Shen H.,
As Merkillow G., Millshina W., Strong R., Weissenbach J., Wang X.,
As Millams S.M., Woodage T., Rorley K.C., Hu D., Yang S., Yao Q.A.,
As Globek R., Myers B.W., Rubin G.M., Weissenbach H., Wang X.,
As Globek R., Myers B.W., Rubin G.M., Weissenbach H., Wang X.,
As Globe R.A., Wassarman D.A., Weinston B. Develler F., Shen H.,
The genome sequence of Drosophila melanogaster "",
REGILA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alternative splicing; Developmental protein; Polymorphism.
DOMAIN 78 280 ASP/GLU-RICH (ACIDIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRINTS: PRO0454; ETSDOMAIN.
PROSITE; PS00345; ETS_DOMAIN_1; 1.
PROSITE; PS00346; ETS_DOMAIN_2; 1.
PROSITE; PS00061; ETS_DOMAIN_3; 1.
Nuclear protein; Transcription regulation; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALA/GLN/SER-RICH.
ARG/LYS-RICH (BASIC).
ETS-DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FlyBase; FBgn0000567; E1p74EF
InterPro; IPR000418; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X15087; CAA33195.1; -. EMBL; M37083; AAA28494.1; -. EMBL; AE003523; AAF49324.1; -. PIR; S04722; S04722. PIR; B34692; B34692.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Ets; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P04002; 1ATF.
TRANSFAC; T00210; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        760
787
27
72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
```

POLY-SER. POLY-ALA.

DOMAIN

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PGCA_MOUSE STANDARD; PRT; 2132 AA.

061282; 064021;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
AGGRECAN CORE PROFEIN PRECURSOR (CARTILAGE-SPECIFIC PROFEOGLYCAN CORE
                                                                                                                                                                                                                                                                                                                        ---KGPL----DILNTQIGPSAFRVEVQADG----THAAIGEKNGLEVSVTLSPQEWSSL 143
                                                                                                                                                                                                                                                                                                                                       QSIDTEGKNRFVFTGGRGGSGHPMV-----TVASDIAEARTRILAKLDPDNHGGRQP 195
                                                                                                                                                                                                                                                                                                                                                                                        : | : : | | | : | | 1
152 EEISNKS------PPVQEDEEESESVASDCREFKVLYNHLRQQQHH--HSP 194
                                                                                                                                                                                                                                                                                                                                                                                                                           196 KDVD-TRSVGVGSASGIDD--GVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATG 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            253 IAQALALTPE-----PDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              304 NG-----NAIP-----SGELXDDIV--EQIAQQAKEAGEVARQQAVE---- 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      284 GGATSAAASAAAAASAVSRRRGRTYSGTESDDSAQCERARMRLKPERKAERSAAYKKSLM 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       344 KRYYTEIPIVKQSTSPAPQQQLQQQHHLQQQQQQPHNGSTF------AGATA 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----LHGAAGNGGSSNGNNAHQQQQPL----AIPQRPLLHNLLSGGAIHNPHHRNYTT 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTT 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  385 ALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSD 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     445 SSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPV------ 494
                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                     SSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHD---- 94
                                                                                                                                                                                                                                                                                             38 SSSLSLSSSSSSSLSSATPTPVAS---PVTPTSPPPAAAAPAEASPPAGAELQEDGQQA 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --TGRLIGTPG-QGIQSTYALLANSGG------LRLGMGGLTSGGETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      195 SSPDKTRST-----LDDVSKILWERKQOLQRSSV-------
                                                                                                                                                                                                                   1 5.1%; Score 144; DB 1; Length 883;
Similarity 20.8%; Pred. No. 0.46;
24; Conservative 68; Mismatches 197; Indels 208;
                                                                                                                                                                               148D5031A18D1409 CRC64;
POLY-ALA.
POLY-GLN.
POLY-GLN.
POLY-GLN.
POLY-SER.
POLY-ALA.
                                                                                     POLY-GLN.
POLY-GLN.
                                                                                                                POLY-ALA.
POLY-SER.
                                                                                                                                         POLY-ALA.
POLY-ALA.
                                                                            GLN
                                                                            POLY-
                                                                                                                                                                               94820 MW;
2999
378
415
4837
524
6630
6642
6661
711
711
717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (Mouse).
                                                                                                                                                                             883 AA;
 289
372
372
4434
434
434
614
614
651
707
712
721
721
867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGC1 OR AGC
                                                                                                                                                                                                                                            Matches 124;
                                                                                                                                                                                                                     Query Match
                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                   Local
                                                                                                                                                                   VARIANT
                                                                                                                                         DOMAIN
DOMAIN
             DOMAIN
DOMAIN
                                                                DOMAIN
                                                                                         DOMAIN
                                                                                                     DOMAIN
                                                                                                                 DOMAIN
                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 4
PGCA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 339
                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

```
this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY).

DOMAIN: TO THE DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINIS OF THE PROTECGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.

DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF CARTILAGE MATRIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFICIENCY (CMD), AN AUTOSOMAL RECESSIVE SYNDROME CHARACTERIZED BY CLEFT PALATE, SHORT LIMBS, TAIL AND SNOUT. MUTATION IN STRAIN CMD CAUSES ABSENCE OF AGGRECAN BY TRUNCATION OF THE PROTEIN (MUTATION IN THE G1 DOMAIN).
                                                                                                                                                                                                                                                                                                                                                                                                      the aggrecan gene."; nat. 17(1994).

Nat. Genet. 7:154-157(1994).

-I- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.

-I- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
-!- SIMILARITY: CONTAINS 4 LINK DOMAINS.
-!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-!- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
-!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
                                                                                                                                                                                                                                                                                                                                        Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus NCBI_TaxID=10090;
                                                                                                                                                                                   "Complete coding sequence, deduced primary structure, chromosomal localization, and structural analysis of murine aggrecan."; Genomics 22:364-371(1994).
                                                                                                         STRAIN-BALB/C; TISSUE-Cartilage;
MEDLINE-9110484 F; PubMed-7806222;
WALCZ E., Deak F; Erhardt P., Coulter S.N., Fueloep C., Horvath
Doege K.J., Glant T.T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00084; sush1; 1.
PROSITE; PS00290; IG_MHC; 1.
PROSITE; PS01241; LINK; 4.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; L07049; AAC37670.1; -
EMBL; S73722; AAB32160.1; -
EMBL; S73721; AAB32160.1; JOINED.
HSSP; P98066; 1TSG.
                                                                                                                                                                                                                                                                                                   STRAIN-129/SV;
MEDLINE-95004579; PubMed-7920633;
                                                                                                                                                                                                                                                                               SEQUENCE OF 211-326 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00193; Xlink; 4.
Pfam; PF00059; lectin_c; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000436; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR000538; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001304; -.
                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                   famada Y.;
```

```
STRDSL-----ETSASGVDVTG--FPSGRGDPETSVSGVGDDF-----SGLP 1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          240 AIAAGLAGLAATGIAQALALTPEPDDPTTTDP------DQAANAAES--- 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64 -RDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAA 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                281 ATKDQLTQEAFKNPENQKVNIDANGNAIPSGE-----LXDDIVEQIAQQAKEAGEV 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123 IGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVAS---DIAEART
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180 RILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..) (POTENTIAL).
..) (POTENTIAL).
..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELL ATTACHMENT SITE (POTENTIAL)

CELL ATTACHMENT SITE (POTENTIAL)

W; 0B2BCDFC6CBDA163 CRCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 129; Conservative 81; Mismatches 241; Indels 178;
Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.0%; Score 140; DB 1; Length 2132; 20.5%; Pred. No. 2.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIM
                                                                                            AGGRECAN CORE PROTEIN. IG-LIKE V-TYPE DOMAIN.
                                                                                                                                                                                      LINK 2.
LINK 3.
LINK 4.
C-TYPE LECTIN.
                                                                  POTENTIAL.
                                                                                                                                                              LINK 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CS-1.
CS-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G2-B'
                                    domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA; 222008
                                    Repeat; Immunoglobulin
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                             DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1054
         KKW
Parteral Series of the Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
```

us-09-189-415a-2.rsp

S

```
Hydrolase;
                                                                                                CHAIN
PROPEP
ACT_SITE
SITE
                                                                                                                                                                                                                                                                    SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YPJA_ECOLI
                                                                                                                                                                                                                                        SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        830
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-90154052; PubMed-2105953;
A MEDLINE-90154052; PubMed-2105953;
Bachovchin W.W., Plaut A.G., Flentke G.R., Lynch M., Kettner C.A.;
Inhibition of 1gA1 proteinases from Neisseria genorrhoeae and
"Inhibition of 1gA1 proteinases from Neisseria and
"Inhibition of 1gA1 proteinase from Neisseria and
"Inhibition of 1gA1 proteinase from Inhibition of 1gA1
"Inhibition of 1gA1 proteinase from Neisseria and
"Inhibition of 1gA1 proteinase from Nei
                                       1250 GEYSGADIGSGPSSGLPDFSGLPSGFPTVSLVDSTLVEVITATISSELEGRGTIGISGSG 1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1354 VSGQPFGSSGVSEETSGIPEISGQPSGTPDTTATSGVTELNELSSGQPDVSGDGSGILFG 1413
                                                                                                                                                                                                                                                                                                                                        -----FFD 1353
                                                                                                                                    ----IGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTHTVVQQQTGGIP-QH 415
                                                                                                                                                                                                                                                                    KVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Proteobacteria; beta subdivision; Neisseriaceae; Neisseria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pohlner J., Halter R., Beyreuther K., Meyer T.F.;
"Gene structure and extracellular secretion of Neisseria gonorrhoeae
                                                                                                                                                                                                                                                                                                                                                                                                     476 VAADP--GYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALL------ANSGGLRLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-WAR-1989 (Rel. 10, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA
                                                                                                                                                                                                                                                                                                           PRT; 1532 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       523 MG---GLTS-GGETA-VSSVNAAPTPGPV 546
ARQQAVESNAQAQQRYEDQHARRQEELQLSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-87115823; PubMed-3027577;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-1989 (Rel. 10, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000710; -.
PRINTS; PR00921; IGASERPTASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; X04835; CAA28538.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nature 325:458-462(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Neisseria gonorrhoeae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A26039; A26039.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S06.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protease."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-1989
01-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACTIVE SITE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IGA_NEIGO
P09790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IGA_NEIGO

ID AZA NEIGO

DT 01-WAI

DE PROTE

GN NEISS

ON                                                                                                                                                                                                                                                                     416
                                                                                                                                364
                                                     g
                                                                                                                            ò
                                                                                                                                                                                      q
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
-----SESSIGNLARVIPRMGRELINDYEEI-------PLE 1164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         980 RPVKPAPSPAANTASQAQKATQTDGAQIAKPONIVV-----APPSPQANQAEEALRQ 1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1032 QAK-AEQVKRQQAAEAEKVARQKDEEAKRKAAEIARQQEEAR------KAAELAAKQ 1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1082 KAEAERKARELAROKAEEAS--------HOANAKPKRRRRAILPRPPAPV 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      470
                                                                                                                                                                                                                                                                                                                                                                              782 TGYVTCNTGNLSDKAL-----NSFDATRINGNVNLNQNAALVLGKAALWGKIQ--- 829
                                                                                                                                                                                                                                                                                                                                                                                                                          94 DKGPLDILNTQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNR 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GQGNSRVSLNQHSKWHLTGDSQVHNLSLADSHIHLNNASDAQSANKYH 877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  212 D--DGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAAT-----GIAQALALTPEPD 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         931 DLFDA-----SSVQDRSRLFVSLANHYVDLGALRYTIKTENGITRLYNPYAGNG 979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   324 QAKEAGEVARQQAVESNAQAQQRYED-----QHARRQEELQLSSGIGYGLSSALIVAGG 377
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                    34 TGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLH 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YPJA_ECOLI STANDARD; PRT; 1569 AA.
PP52143; PP6610; P77017; P77019;
01-0CT-1996 (Rel. 34, Created)
01-NOY-1997 (Rel. 35, Last sequence update)
01-0CT-2000 (Rel. 40, Last annotation update)
11-YPOTHETICAL 98.4 KDA PROTEIN IN ALPA-GABD INTERGENIC REGION (F949).
YPJA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINE-97426617; bubMed=9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      154 FVFTGGRGSG--HPMVTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   265 DPTTTDPDQAANAAESATKDQLTQEA-FKNPENQKVNIDANGNAIPSGELXDDIVEQIAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       378 IGAGVTTALHRRNOPAEOTTTTTTTTTTTTVQQQTGGIPQHKVALMPQERRR------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      427 FS----DRRDSQGSVASTHWSDSS------SEVVNPYAEVGGARNSLSAHQPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; gamma subdivision; Enterobacterlaceae;
Escherichia.
                                                                                                                                                                                                                                                                                            183; Indels 140;
                                                             SERINE ENDOPEPTIDASE.
                                                                                                                                                                                                                                                    Length 1532;
  Zymogen; Autocatalytic cleavage;
                                                                                                                       CLEAVAGE (AUTO-).
CLEAVAGE (AUTO-).
CLEAVAGE (AUTO-).
W; 68FF4112BD22F40D CRC64;
                                                                                                                                                                                                                                                    DB 1;
                                                           IGA-SPECIFIC SE
HELPER PEPTIDE.
                                                                                                                                                                                                                                     Score 137; DB
                                                                                                                                                                                                                                                                      ; Pred. No. 2.3; 60; Mismatches
                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                    4.9%;
Serine protease;
                                                                                                                                                                                    AA; 168976
                                                                                                                                                                                                                                                                        Local Similarity 21.0
nes 102; Conservative
                                                         986
1532
278
987
1019
                      Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              || | :|:
1125 FSLDDYDAKDN-
                                                                                                   278
986
1018
11121
1532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1165 ELEDE 1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=562;
                        Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          471 HIYDE 475
```

```
PGCA_RAT
P07897;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This
                                                                                                                                                                                                       7
                                                                                                                                                                                                                               PGCA_RAT
                                                                                                                                                                                                                                                    g
                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                              this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1008 DLOVAGGTAIVYAGTLAD-----ASVSGATGSLSLMTPRDNV------TPVKLE 1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  956 LVLAGTEARDSTVGKGGAMQNLGQDSATKVN-----SGG--QYTLGRSKDEFQALARAE 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               896 GGTLTVNDNGKATDIVQNSGAALQTSTANGIEISGTHQYGTFSISGNLATNMLLENGGNL 955
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        372 LIVAGGIGAGVITALHRRNQPAEQTITITITITYVQQQTGGIPQHKVALMPQERRRFSDRR 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            289 EAFRNPENQKVNIDANGNAIPSGELXDDIVEQIA------QQAKEAGEVARQQ--- 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------ATGTTLTGGE---QIVEGVANETTINDGGIQTVSANGEAIKTKINE 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77 AAATSETCLLGGFEVLHDKGPLDILNTQIGP---SAFRVEVQADGTHAAIGEKNGLEVSV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        629 TVVTGSRAV---DIIINANGKMDVYGKDVGTVLNSAGTQTIYASATSDKANIKGGKQTVY 685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    134 TLSPQE--WSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTRILAKLDPD--- 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 189 -----NHGGRQPKDV-----DTRSVGVGSASGI---DDGVVSETHTSTTNSSVRS 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DPKFWVSVGAIAAGLAGLAATGI--AQALALTPEPDDPTTTDPDQAANAAESATKDQLTQ 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             D-----GVAFSIEQGAANNILLANGGVLTVESD--TSSDKTQVNMGGREIVKTKAT- 852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   579 AKHVEQQSGGALIASTISGILIEGINSYGDAFY---IRNSEAKNV-----VLENAGSL 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19 APPLPSQTDGAARGGT--GHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRL 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               432 DSQ--GSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFS
                                                                                                                   Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T., Ikemoto K., Inada T., Isono K., Isono S., Itoh T., Kanai K., Kasal H., Kashimoto K., Kimura S., Kitagawa M., Kitakawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N., Sampel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., A., Marana M., Marana M., Marana M., Marana M., Mana M., Mana M., Manamoto Y., Yano M., Mana M., Mana M., Mana M., Manamoto Y., Yano M., Mana M., Manamoto Y., Manamoto M., Mana M., Manamoto 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.9%; Score 136; DB 1; Length 1569;
20.9%; Pred. No. 2.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20.9%; Pred. No. 2.7;
Live 69; Mismatches 259; Indels
         "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                              Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
--- SUBCELLULAR LOCATION: OUTER MEMBRANE (POTENTIAL).
--- SIMILARITY: STRONG, TO BORDETELLA PERTACTIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0D006A9A32154596 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---AVESNAQAQQRYEDQHARRQEE----LQLSSGIGYGLSS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL, D90889; BAA16514.1; ALT_INIT.
EMBL, D90890; BAA16518.1; ALT_INIT.
ECOGENE: EG13213; PpjA.
Hypothetical protein; Outer membrane.
SEQUENCE 1569 AA; 162774 MW; ODOG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U36840; AAA79815.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AE000350; AAC75695.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 20.9% Matches 127; Conservative
                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
SIMILARITY: CONTAINS 1 IMMUNOSLOBULIN-LIKE V-TYPE DOMAIN.
SIMILARITY: CONTAINS 1 LINK DOMAIN.
SIMILARITY: CONTAINS 1 LINK DOMAIN.
SIMILARITY: CONTAINS 1 LINK LECTIN FAMILY DOMAIN.
SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.
SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, MOTIFES, G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1856-2124 FROM N.A.

MEDLINE-86250698; PubMed=2424893;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
J. Biol. Chem. 261:8108-8111(1986).
J. Biol. Chem. 261:8111(1986).
J. Biol. Chem. 261:811(1986).
J. Biol. C
                                                                                         1051 GAVRITDSATLTLGNGVDTTLADLTAASRGSVWLNSNNSCAGTSNCEYRVNSLLLNDGDV 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-AUG-1988 (Rel. 08, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
AGGREGAN CORE PROTEIN PRECURSOR (CARTILAGE-SPECIFIC PROTEOGLYCAN CORE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Mammalia, Butheria, Rodentia, Sciurognathi, Muridae, Murinae, Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIINE-88087070; PubMed-3693370;
Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
Doege K., Saraki M., Horigan E., Hassell J.R., Yamada Y.;
Complete primary structure of the rat cartilage proteoglycan
protein deduced from cDNA clones.";
J. Biol. Chem. 262:17757-17767(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     between the Swiss Institute of Bioinformatics and the the European Bioinformatics Institute. There are no re-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada
J. Biol. Chem. 263:10040-10040(1988).
490 GSGPVTGRLIGTPGQGIQSTYALL--ANSGGLRLGMGGLTSG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M13518; AAA41836.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                         1111 YLSAQTAAP 1119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                              533 AVSSVNAAP 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REVISION TO 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OR AGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AND G3
```

440

```
----DVSGLPSGKEGLETSASGIEDISVFPTEAEGLETSASGGYVSGIPSGEDGTETSTS 1117
                                                                                                                                                                                        1118 GVEGVSGLPSGGEGLETSASGVEDLGLPTRDSL-----ETSASGVDVTG--YPSGRE 1167
                                                                                                                                                                                                                                                          1278 STLVEVITATTASELEGRGTISVSGSGEESGPPLSEL------DSSADISGLPSGT 1327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----FFDVSGQPFGS----SGTGEGTSGIPE 1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLPSGEEGSETST----SGIEDISVLPTGESPETSASGVGDLSGLPSGGESLETSASGVE 1021
                                 176 EARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSET---HTSTTNSSVRSDP 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GLRLGMG---GLTS-GGETA-VSSVNAA 540
                                                                ----SASGIEDITVLPTGRENLETSASGVE---
                                                                                                                                                                    ---DQAANAAES---ATKDQLTQEAFKNPENQKVNIDANGNAIPSGE-
                                                                                                                                                                                                                                     -----LXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSG-
                                                                                                                                                                                                                                                                                                      ----IGYGLSSALIVAGGIGAGVTTALHRR
                                                                                                                                                                                                                                                                                                                                                                         390 NOPAEQTTTTTTTTTTTVVQQQT-----GGIPQHKVALMPQERRFSDRRDSQGSVAST
                                                                                                                                                                                                                                                                                                                                                                                                                                         441 HWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIG
                                                                                                   233 KFWVSVGAIAAGLAGL--AATGI-----AQALALT-----PEPDDPTTTDP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Euteleostomi;
Sus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Submaxillary gland;
MEDLINE-88087170; PubMed-2826455;
Timpte C.S., Eckhardt A.E., Abernethy J.L., Hill R.L.;
"Porcine submaxillary gland apomucin contains tandemly repeated, identical sequences of 81 residues.";
J. Biol. Chem. 263:1081-1088(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
TISSUE-Submaxillary gland;
MEDILINE-91236743; PubMed-2033060;
Eckhardt A.E., Timpte C.S., Abernethy J.L., Zhao Y., Hill R.L.;
"Porcine submaxillary mucin contains a cystine-rich,
carboxyl-terminal domain in addition to a highly repetitive,
glycosylated domain.";
J. Biol. Chem. 266:9678-9686(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Submaxillary gland;
MEDLINE-87280230; PubMed-3611111;
Eckhardt A.E., Timpte C.S., Abernethy J.L., Toumadje A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0CT-1989 (Rel. 12, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
01-0CT-2000 (Rel. 40, Last annotation update)
APOMUCIN (MUCIN CORE PROTEIN) (FRAGMENT).
Sus scrofa (Pig).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Cetartiodactyla; Sulna; Suldae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                         DV-TQL-----PTERGGLET---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1-503 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1328 ELSGQTSGSLDVSGETSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          501 TPGQGIQS--TYALLANSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 45-80.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |:- ||
1429 PSGFPV 1434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9823;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             541 PTPGPV 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APMU_PIG
P12021;
                                                                  1022
                                                                                                                                   1062
                                                                                                                                                                                                                                     313
                                                                                                                                                                                                                                                                                                      364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APMU_PIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                               g
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                   Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                               g
                                                                                                  ŏ
                                                                                                                                                                   ò
                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DIPGLPTNPSRLAAATSETCLLGGFEVL-HDKGPLDILNTQIGPSAFRVEVQADGTHAAI 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GEKNGLEVSVTLSPQEWSSLQSIDT -- EGKNRFVFTGGRG-----GSGHPMVTVASDIA 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PPAPPLPSQTD----GAARGGTGHL-----ISSTGALGSRSLFSPLRNSMADSVDSR 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2124;
                                                                                                                          Pfam; PP00193; X1ink; 4.
Pfam; PP00059; lectin_c; 1.
Pfam; PP00059; lectin_c; 1.
PROSITE; PS00290; IG_MHC; 1.
PROSITE; PS01241; LINK; 4.
PROSITE; PS50041; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                   POTENTIAL.
AGGRECAN CORE PROTEIN.
IG-LIKE V-TYPE DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 135.5; DB 1;
Pred. No. 4.2;
7; Mismatches 204;
                                                                                                                                                                                                                                                                                                                                                                                                                                  G1-A.

G1-B.

G1-B.

G2-B.

G2-B.

G2-B.

CS-1.

CS-2.

G3-B.

G3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      880 PTVDRLPSGGESLEGSASASGTGDLSGLPSGGEITETSASG-
                                                                                                                                                                                                                                                                                                                                    LINK 1.
LINK 2.
LINK 3.
LINK 4.
C-TYPE LECTIN.
SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 , 77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 221117
AAA21000.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                    ; IPR000436; -.; IPR000495; -.
                                                                                                                   IPR001304; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1925
2034
2026
2084
2097
                                               HSSP; P98066; 1TSG.
InterPro; IPR000436;
               A23835; A23835.
A28452; A28452.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
J03485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 138;
                                                                                  InterPro;
                                                                                                   InterPro;
                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           124
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
```

```
RLG------MGGLTSGGETAVSSVNAAP 541
                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local 8
                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                           CARBOHYD
        CARBOHYD
CARBOHYD
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                           CARBOHYD
                                                                                                                                          CARBOHYD
                                                                                                                                                               CARBOHYD
                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                   CARBOHYD
                                           CARBOHYD
                                                        CARBOHYD
                                                                                          CARBOHYD
                                                                                                    CARBOHYD
                                                                                                                  CARBOHYD
                                                                                                                                                     CARBOHYD
                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OY
OY
Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                     PTM: EXTENSIVELY O-LINKED GLYCOSYLATED ON SER AND THR RESIDUES OF THE REPEAT UNITS. HIGHEST GLYCOSYLATION APPEARS TO OCCUR ON SER RESIDUES WHICH HAVE GLY AT POSITIONS AT ±2 OR -2 FROM THE GLYCOSYLATION SITE OR, WHERE GLY IS THE PENULTIMATE RESIDUE. THE PRESENCE OF PROLINE (USUALLY AT POSITION +3 OR -3) APPEARS TO ALSO ENHANCE GLYCOSYLATION.
                                                                                                                                                                                                                                                                                                                                               -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
                                                                             Gerken T.A., Owens C.L., Passumarthy M.;
Gerken T.A., Owens C.L., Passumarthy M.;
Determination of the site-specific O-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:galnac transferase peptide binding site:";
J. Biol. Chem. 272:9709-9719(1997).
-1- FUNCTION: APOMUCIN IS PARY OF WOIN, THE MAJOR GLYCOPROTEIN SYNTHESIZED AND SECRETED BY WICOUS CELLS OF THE SUBMAXILLARY GLAND. ITS HIGHLY VISCOUS AQUEOUS SOLUTIONS SERVE TO LUBRICATE THE ORAL CAVITY AND TO PROTECT IT FROM THE EXTERNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
                                                                                                                                                                                                                  MULTIMERIC MUCIN STRUCTURE.
1- SUBCELLULAR LOCATION: SECRETED.
1- TISSUE SPECIFICITY: SUBMAXILLARY MUCOSAE.
1- TISSUE SPECIFICITY: SUBMAXILLARY MUCOSAE.
1- DOMAIN: CONTAINS TANDEMLY REPEATED, IDENTICAL SEQUENCES OF 81
Johnson W.C. Jr., Hill R.L.; "Structural properties of porcine submaxillary gland apomucin."; J. Biol. Chem. 262:11339-11344(1987).
                                                                                                                                                                                                    SUBUNIT: INTERMOLECULAR DISULFIDE BONDS COULD HELP MAINTAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AA TANDEM REPEATS
                                              CARBOHYDRATE-LINKAGE SITES, AND SEQUENCE OF 45-125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTCK.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

O-LINKED (GALNAC.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6 (INCOMPLETE)
VWFC.
                                                              TISSUE-Submaxillary gland;
MEDLINE-97248516; PubMed-9092502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M61883; AAA30990.1; --
EMBL; M21174; AAA30990.1; --
PIR; A40009; A40009.
InterPro; IPR001559; --
InterPro; IPR001007; --
Pfam; PF00007; Cys_Knot; 1.
PROSITE; PS01128; CTCK_1; 1.
PROSITE; PS01225; CTCK_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1146
1109
1123
1139
1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46
50
57
57
61
66
67
                                                                                                                                                                                               ENVIRONMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
```

```
249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SLSAHQPEEHIYDEVAADPGYSVIQNFSGSGP-VTGRLIGTP-GQGIQSTYALLANSGGL 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGAGPGITASSVGVTETARPSVAGSGTTGTVSGASGSTGSSSGSPGATGASIGQPETSRI 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSSEVVNPYAEVGGARN 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                250 VAGSGTTGTVSGASGSTGSSSGSPGATGASIGOPETSRISVAGSSGAPAVSSGASQAAGT 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            246 AG-----LAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKV 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             300 NI--DANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEE 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142 SLQSIDTEGKNRFVFTGGRGG------SGHPMVTVASDIAEARTRILAKLDPD 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91 SGTIGIVSGASGSIGSSSGSPGATGASIGOP-EISRISVAGSSGAPAVSSGASO-AAGIS 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82 ETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWS 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  149 ------GAGPGTTASSVGVTETARPSVAGSGTTGTVSGASGSTGSSGSPGATG 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 SQTDGAARGGTGHLISSTGALGS--RSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATS 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   189 NHGGRQPKDVDTRSVGVGSASGIDDGV---VSETHTSTTNSSVRSDPKFWVSVGAIAAGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :: :: |: || || || || SVAGSSGAPAVSSG--------SQA--AGTSEATTSIEGAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 358 LQLSSGIGYGLSSALIVAG-----GIGAGVTTALHRRN-QP----AEQTTTT----TT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---TSGVGF-KTEATTFPGENETTRVGIATGTTGIVSRKTLEPGSYNTEATTSIGRSGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HT-----DLPGGTTIVLPGFSH--SSQSSKPGSSVTTPGSPESGSETGTSGEFSTTVI
                                                                                              (MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
(MUCIN TYPE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1150;
                                                       (MUCIN 1
(MUCIN 1
(MUCIN 1
(MUCIN 1
(MUCIN 1
(MUCIN 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3CB68B5D29DD7F5A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60; Mismatches 260;
                                                                                                                                                                                    (GALNAC.
(GALNAC.
(GALNAC.
(GALNAC.
(GALNAC.
(GALNAC.
                                                                                                                                                                                                                                                                                                                                   (GALNAC.
(GALNAC.
(GALNAC.
                                                                                                                                                                                                                                                                                                                                                                                                     (GALNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GALNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.8%; Score 134.5; 22.1%; Pred. No. 2.2
O-LINKED
O-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109615 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 126; Conservative
  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
```

σ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Awram P., Smit J.K.;
"The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (type I) secretion apparatus.";
J. Bacteriol. 180:3362-3069(1998).
-i. FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
-i. SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S-LAYER WITH HEXACONAL SYMMETRY.
MISCELLANEOUS: THE CALCIUM BINDING OF THIS PROTEIN MAY MEDIATE
SURFACE ATTACHMENT BY IONIC BRIDGING BETWEEN THE S-LAYER AND A
SPECIFIC MEMBRANE-ASSOCIATED OLIGOSACCHARIDE.
                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A., AND SEQUENCE OF 1-21; 646-660 AND 1021-1025. STAIRA-MACC 19089 / CB15. MEDIJNE-93007499; PubMed-1393820; Gilchrist A., Fisher J.A., Smit J.K.; Gilchrist A., Fisher Jaysis of the gene encoding the Caulobacter crescentus paracrystalline surface layer protein."; cascentus paracrystalline surface layer protein."; can. J. Microbiol. 38:193-202(1992).
                                                                                                                                                                                                                                                                    Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-ATCC 19089 / CE15;
MEDLINE-89008089; PubMed=3049545;
Fisher J.A., Smit J.K., Agabian N.;
"Transcriptional analysis of the major surface array gene of Caulobacter crescentus.";
                                                                                                                                                                      01-OCT-1994 (Rel. 30, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
S-LAYER PROTEIN (PARACRYSTALLINE SURFACE LAYER PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1025 AA; 98209 MW; AFC8B519820B1A5F CRC64;
                                                                                                               PRT; 1025 AA
Cell wall; S-layer; Calcium-binding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00353; hemolysinCabind; 1. PRINTS; PR00313; CABNDNGRPT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteriol. 170:4706-4713(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=98292737; PubMed=9620954;
                                                                                                           SLAP_CAUCR STANDARD; P35828; Q46015; O1-JUN-1994 (Rel. 29, Created) 01-CCT-1994 (Rel. 30, Last seq 30-MAY-2000 (Rel. 39, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AF062345; AAC38665.1; -.
                                                                                                                                                                                                                                                    Caulobacter crescentus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001343;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                         NCBI_TaxID=76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                       g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Haloferax volcanii.";
J. Bacteriol. 172:7111-7118(1990)
-!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
-!- FUNCTION: THE S-LAYER US A PRACATOR THE SURFACE OF THE CELL.
-!- SUBCELLULAR LOCATION: CELL WALL. THIS ARCHAEA IS COVERED BY A
S-LAYER WITH HEXAGONAL SYMMETRY.
-!- PTM: PROTEIN IS CONTAINS O-LINKED GLYCANS WHICH CONSIST OF GLC-GAL
                                                                                                                                                                                                                                                                                                                                                                            248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     519
229 TDNAAGVNLFTAYPSSGVSGSTL-----SLTTGTDT--LTGTANNDTFVAGEVAGAAT 279
                                                                          --TQIGPSAFRVEVQADGTHAAIGEKNGLEVSVT 134
                                                                                                                                                 280 LTVGDTLSGGAGTDVLNWVQAAAVTALPTGVTISGIETMNVTSGAAITLNTSSGVTGLTA 339
                                                                                                                                                                                                                       LSPOEWSSLOSIDT-EGKNRFVFTGGRGGS----GHPMVTVASDIAEARTRILAKLDPD 188
                                                                                                                                                                                                                                                                                                                                                                                                                           | :| || : | || : | || : | || : | || : | : | : || : | : || : | : | : || : | : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             249 AATGIAQALALIPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAI 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   360 LSSGIGYGLSSALIVAGGIGAGŸTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVAL 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               604 AAALTGITVTNSVGATLGAELATG-LVFTGGAGRDSILLGATTKAIVMGAGDDTVTVSSA 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Halobacterium volcanii (Haloferax volcanii).
Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloferax.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           340 INTUTSGAAQTVTAGAGQNLTATTAAQAANNVAVDGRANVTVAS-----TGVTS----
                                                                                                                                                                                                                                                                                                                                                                            NHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              309 PSGELXD--DIVEQIAQQAKEAGEVAR-----QQAVESNAQAQQRYEDQHARRQEELQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     420 MPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEV----GGARNSLSAHQPEEHIYDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         476 VAADPGYSVIQNFS---GSGPVTGRLIGTPGQGIQS-----TYALLANSG-----GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --SGTGTSL-----GIGRGALTA----TPTANTLTLNVNGLT--TTGAITDSEAA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDITNE-91072265; PubMed-212862;
Sumper M., Berg E., Mengele R., Strobel I.;
"Primary structure and glycosylation of the S-layer protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISACCHARIDES.
SIMILARITY: TO H.HALOBIUM CELL SURFACE GLYCOPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
01-CTT-2000 (Rel. 40, Last annotation update)
CELL SURFACE GIXCOPROTEIN PRECURSOR (S-LAYER PROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       424 TAVTVAQTAG---NAVNTTLTQADVTVTGNSSTTAVTVTQTA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          520 RLGMGGLTSGG---ETAVSSVNAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    663 TLGAGGSVNGGDGTDVLVANVNGS 686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                          LGGFEVLHDKGPLDILN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=2246;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CSG_HALVO
P25062;
                                                                          98
                                                                                                                                                                                                                                                                                                                                                                                                                                                      389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CSG_HALVO
                                                                                                                                                 g
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
```

27;

Gaps

Indels 155;

64; Mismatches 222;

Matches 123; Conservative

ò

Query Match Best Local S

Length 1025;

1, 4.7%; Score 132.5; DB 1; Similarity 21.8%; Pred. No. 2.5;

26 IDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCL 85

```
RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MAZ4_SCHCO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
ò
                                                                                                                                                                                                                                                                                                                                                                                                                       26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             384 TALHR---RNOPAEQTTTTTHT----VVQQQTGGIPQHKVALMPQERRRFSD 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             430 RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFS 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----TGTVDINGTASGANSVLVIFVDERGNVN------YQEVSVDSDGTYDEDDI 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             274 AANAA--ESATKD-QLTQE----AFKNPENQKV---NIDANGNAIPSGELXDDIVEQIAQ 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    364 SATAAVGQDATNDITLTIEEGGTTLSSPTGQYVVGSEVDINGTATSS-----DSVAI 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           324 QAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVT 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           116 YVRDDGD---WQLLEIGGDNEISVDSDDTFEEEDIALSGLSGDG-SSILSLTGTYRIGVI 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            154 FVFTGGRGGSGHPMYTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDD 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      262 YTVTNGIDGNEHVVAMDLSDLQNDATT------EQAKEV-FRNIGDTSEVGIAN 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             214 GVVSETHTSTINSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQ 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    309 SSATNTSGSSTGPTVET----ADIAYAVVEIDGASAVGGIETQYLDDSEVDLEVYDAGV 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56 SMADSVD-----SRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPS 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 PIGNLGNNVNGNHLIPPAPPL-PSQTDG---AARGG--TGHLISSTGALGSRSLFSPLRN 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472 DASDADVGGDGSVDDSLITSEFTSGVSSSNSIRVTDQALIGQFTTINGQVAPVE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    160 ESAEKVEVTVEDPSGIDI------TNE--VLSGIDIFVDDGSIGSISSIGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109 AFRVEVQADGTHAAIGEKNGLE------VSVTLSPQEWSSLQSIDTEGKN----R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GIDMSDQDAGEYTIILE--GAEDLDFGDATETMTLTISSQDEIGIE-LDSESVTQGTDVQ
                                                                                                                                                                                                                                                                                                                                                                                                                   81; Mismatches 243; Indels 132;
                                                                                                                                                                                                                             LINKED (GLC...) (POTENTIAL)
LINKED (GLC...) (POTENTIAL)
-LINKED (GLC...) (POTENTIAL)
-LINKED (GLC...) (POTENTIAL)
B208E96283699974 CRC64;
                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                            Query Match 4.7%; Score 132; DB 1; Length 827; Best Local Similarity 21.4%; Pred. No. 2;
                                                                     Cell wall; S-layer; Signal.
                                                                                                              SURFACE GLYCOPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSG 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   618 AA
                                                                                                                                                                           (GIC.
                                                                                                                                                                           N-LINKED (N-LINKED (N-LINK
                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Drosophila virilis (Fruit fly)
                                                                                                                                                                                                                                                                                                                          ΜŇ,
                                                                         Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REGULATORY PROTEIN ZESTE.
                                                                                                                                                                                                                                                                                                                        85188
        EMBL; M62816; AAA72996.1;
PIR; A37849; A37849.
GlycoSuiteDB; P25062; -.
                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 21.4*
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                  35
804
804
47
1117
308
313
404
532
                                                                           Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZEST_DROVI
Q24762;
                                                                                                                  CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZEST_DROVI
ID ZEST_D
AC Q24762
DT 15-JUI
DT 15-JUI
DE REGULA
GN Z.
                                                                                                 SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15;
                                                                                                                                                                                                                                                              FlyBase; FBGN0013149; Dvlr'A:
DNA-binding; Transcription regulation; Nuclear protein.
DNA-binding; Transcription regulation; WITH ZESTE LOCUS.
DNA_BIND 66 146 GLN/ALA-RICH (OPA-REPERT INVOLVED IN GLN/ALA-RICH (OPA-REPERT INVOLVED IN ATMSCRIPTIONAL ACTIVATION OR REPRESSION AT DIFFERENT TARGET LOCI) (POTENTIAL).
AT DIFFERENT TARGET LOCI) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        377 GIGAGVTT---ALH----RRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSD 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                217 SETHTSTINSSVRSDPKFWVSV-----GAIAAGLAGLAATGIAQALALTPEPDDP 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     120 SHPHVSAMR-NIKQIQKFWLNSRLRKQYPYRDGSASGGSAGLGKVGTVSASAQQQQQQQQQ 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                267 ITTDPDQAA--NAAESATKDQLTQEAFK-NPENQKVN-IDANGNAIPSGELXDDIVEQIA 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              179 QOQQQQQQHHDSVKVEPEXQISPEASEHNPQGEPFDEIEMDGNDV--SEWEDDPLE--A 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           323 QQAKEAGEVARQQ-AVESNAQAQQRYEDQHA-RRQEELQLSSGIGYGLSSAL----IVAG 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235 QQQQQQQQQQQQAAAQAABAQQQQQQQQSAVAEMQKLQVSAAVAAANASMLNTHRINVD 294
                                                                                                                                                                     MEDLINE-92123185; PubMed=1732733;
Chen J.D., Chan C.S., Pirrotta V.;
"Conserved DNA binding and self-association domains of the Drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 295 SISAEKLTLNDLLHFKPARHDEIILQIKHPTDATATO-----IHTIPAQPQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.7%; Score 131.5; DB 1; Length 618; 22.6%; Pred. No. 1.5;
Live 57; Mismatches 116; Indels 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-007-1994 (Rel. 30, Created)
01-007-1994 (Rel. 30, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
MATING-TYPE PROTEIN A-ALPHA Z4.
Schizophyllum commune (Bracket fungus).
Bukaryota; Fungi; Basidiomycota; Hymenomycetes; Stereales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            430 RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FBqn0013149; Dvir\z.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M76700; AAA29052.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                           [1]
SEQUENCE FROM N.A.
                                                                                                    NCBI_TaxID=7244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MAZ4_SCHCO
P37938;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99
```

```
435 GSVASTHWSDSSSEVVNPYAEV 456
                                                                                                                                                                                                                                 NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             983
                                                                                                  ALS1_CANAL
P46590;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                          RESULT 13
ALS1_CANAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                ŏ
                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PVKIAGAKRGRNDDEEVSPLAKKPR-IFSPPVRPRPQAIRVSLPSPAPSSRGSTPTSPVS 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73 PS----RLAAATSETCLLGGFEV-LHDKGPLDILNTQIGPSAFRVEVQADGTH-AAIGEK 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            127 NGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTRILAKLD 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                187 PDNHGGRQPKD------VD------TRSVGVGSASGIDDGVVSETHTST 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----OAPRDLTAATKSSAGCSVDAVPLPGKSRSLTRSPSISSIS-----SACST 675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TNSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATK 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ::| :| | | | : : | ::| SSSGSDTDSLFSVT-----SDATDIT-EPDEATTADETTTQSTSASSSR 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DQLTQEAFKNPENQKVNID-----ANGNAIPSGEL-----XDDIVEQIA 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DITSQQKRMPPLSIDPRFDPALWSKYDLSPPADGRLHPSDGLRPSAFVPTKLDVRVANLA 778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAG---GIG 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QNPARHWSASKRSPTRASHAAAPIVSYHHATGSIASPAQVAFGEGQLTSVLATGQKAGNA 838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQ-----HKVALMPQERRFFSDRRDSQ 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63; Mismatches 205; Indels 121; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23 PSQTDGAARGGTGHLISSTGALGSRSLFSP-----LRNSMADSVDSR--DIPGLPTN 72
                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00027: HOMEOBOX_1; FALSE_NEG.
PROSITE; PS50071; HOMEOBOX_2; FALSE_NEG.
DRASIND HOMEOBOX_2; FALSE_NEG.
DNA_BIND HOMEOBOX (TALB-TYPE).
DNA_BIND HOMEOBOX (TALB-TYPE).
DOMAIN 370 402 ASP/GLU-RICH (ACIDIC).
DOMAIN 436 469 ASP/GLU-RICH (ACIDIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SSPSVS---SPPSVSVSLPLPSRG----VPSGGIKVTGDPTPWVNWDL-EAHT-----
                                                                                                            dissillar multiallelic homeodomain proteins.";

Proc. Natl. Acad. Sci. U.S. 89:7169-7173(1992).

-I- FUNCTION: SPECIFIES A-ALPHA-4 MATING-TYPE. MAY REGULATE THE EXPRESSION OF GENES SPECIFIC TO THE HOMOKARYOTIC CELL TYPE.

-I- SUBCELDULAR LOCATION: NOCLEAR.

-I- DEVELOPMENTAL STAGE: EXPRESSED CONTITUTIVELY IN HOMOKARYONS.

-I- SIMILARITY: BELONGS TO THE TALE/M-ATYP FAMILY OF HOMEOBOX
                                                                                                  "The A alpha mating locus of Schizophyllum commune encodes two
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 130.5; DB 1; Length 940;
                                                                          Stankis M.M., Specht C.A., Yang H., Glasson L., Ullrich R.C.,
Novotny C.P.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  4B99CBAEDB39621E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 2.9;
                                                             MEDLINE=92357793; PubMed=1353886;
Schizophyllaceae; Schizophyllum
                                                                                                                                                                                                                                                                                                                                                                                                                                                  940 AA; 101856 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.78;
                                                                                                                                                                                                                                                                                                                                  EMBL; M97181; AAB01372.1; -. PIR; D37271; D37271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 113; Conservative
                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001356; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
             NCBI_TaxID=5334;
                                                                                                                                                                                                       PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                        Novotny
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9/9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           380
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                  Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;
"Candida albicans ALS1: domains related to a Saccharomyces cerevisiae sexual agglutinin separated by a repeating motif.";
Mol. Microbiol. 15:39-54(1995).
-!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
-!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-THR.
POLY-SER.
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Candida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [-10.
X 26 AA APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGGLUTININ-LIKE PROTEIN 1.
10 X 36 AA TANDEM REPEATS.
                                                                                                                                                                  32, Last sequence update)
40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Repeat; Signal. POTENTIAL.
                                                                                                       PRT; 1260 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-THR
                                                                                                                                                                                                              AGGLUTININ-LIKE PROTEIN 1 PRECURSOR
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN=ATCC 11651 / B792;
MEDLINE=95272392; PubMed=7752895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           899 LSKSSP--VDSADTVRTRLAEI 918
                                                                                                                                            01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last seq
01-OCT-2000 (Rel. 40, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L25902; AAC41649.2; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cell adhesion; Glycoprotein;
                                                                                                                                                                                                                                                      Candida albicans (Yeast).
                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    684
720
720
792
11152
11152
1152
404
404
455
491
523
633
633
```

```
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1065 -----EEDNKSGSH----ESASTSLKPSMGENSGLTTSTEIEATTTSPTEAPSPAVSSG 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1115 T------DVTTEPTDTREQPTTLSTTSKTNSESVATTQATNENGGKSPS 1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1011 EPTNTREOPTTLSTISNSITEDITISOPTGDN-----GDNTSSINPVPTVATSTLASAS- 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidee; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                               954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66 IPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRV--EVQADGTHAAI 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124 GEKNGLEVSVTLSPQEWSSLQSIDTE---GKNRFVFTGGRGGSGHPMVTVA-SDIAEART 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEVELOPMENTAL STAGE: DURING EARLY NEUROGENESIS MAM PRODUCTS ARE UBIQUITOUSLY LOCATED. DURING LATER STAGES THEY ACCUMULATE IN THE CENTRAL NERVOUS SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MISCELLANEOUS: THE PROTEIN HAS MANY AA HOMOPOLYMERIC DOMAINS: 21 POLY-GLN RUNS (FROM 5 TO 16 AA IN LENGTH), 4 POLY-GLY (6 TO 10 AA), 3 POLY-ASN (3 X 5 AA), 1 POLY-ALA (10 AA) AND 1 POLY-THR (5 AA) RUNS.
                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                           16 IPPAPPLP---SQTDGAARGGTGHLISS-----TGALGSRSLFSPLRNSMADSVDSRD 65
                                                                                                                                                                                                                                                                                                                                                                                                                            180 RILLAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240 AIAAGLAGLAATGIAQALALTPEPDD----PTTTDPDQAANAAESATKDQLTQEAFKNPE
                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC. . .) (POTENTIAL). (GLCNAC. . .) (POTENTIAL). (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          296 NQ-----KVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESN 340
                                                                                                                                                                                                                                                                                        DB 1; Length 1260;
                                                                                                                                                                                                                                                                                  Query Match
4.6%; Score 130; DB 1; Length 12
Best Local Similarity 20.6%; Pred. No. 4.4;
Matches 73; Conservative 48; Mismatches 167; Indels
                                                                                                                                                                                                        763D1063A2354C24 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-91065516; PubMed-1701150;
Smoller D., Friedel C., Schmid A., Bettler D., Lam L.,
Yedvobnick B.;
                                                    (GLCNAC. . .)
(GLCNAC. . .)
(GLCNAC. . .)
                              GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAR-1992 (Rel. 21, Last annotation update)
N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                           AA; 132641 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NEUROGENIC PROTEIN MASTERMIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
     687
723
820
886
918
973
1045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-CANTON-S;
  687
723
820
886
918.
973
1045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAM_DROME
P21519;
                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MAM_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
     FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                    임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----VGVGVGVGRGPNPGQQQQQPNQNMSNANVPSDGFSLSQSQSMNFNQQQQQAAAQQ 1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1047 QOV---OPNMRQRQTQAQAAAAAAAAAAQAQAAANASGP-----NVPLMQQPOV---- 1092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1093 GVGVGVGVGVG-----GVGVGNGGVVGGPGSGGPNNGA-MNQMGGPMGGMPGMQMGGPMNP 1146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 415 HKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYD 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           933 SGDMKSNVSVAAQQGVFFSQQQAQQQQQQQQQQGGTNGPNPQQQQQQQPHGGNAGGGVGVG- 991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                314 XDDIVEQIAQQAKEAGEVARQQAVESNAQAQRYEDQHARRQEELQ----LSSGIGYGL 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          369 SSALIVAGGIGAGVTTALHRRNQPAEQTTT------TTTHTVVQQQTGGIPQ 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     262 EPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNID-----ANGNAIPSGE---L 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
SIMILARITY: TO OTHER NUCLEAR PROTEINS OF DROSOPHILA, TO CERTAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          873 QONNPNTGPGGNTPNAPQQQQQQSTTTTLQMKQTQQLHISQQGGGAQGIQVSAGQHLHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72;
                                                                                                                                                                                                                                                                                                                                                                                       5 X 2 AA TANDEM REPEATS OF G-V.
ALA-RICH.

8 X 2 AA TANDEM REPEATS OF V-G.
7 X 2 AA TANDEM REPEATS OF G-V.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.6%; Score 130; DB 1; Length 1596; 20.6%; Pred. No. 5.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            50; Mismatches 140; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ASP/GLU-RICH (ACIDIC).
W; B944D86EF359D605 CRC64;
                                                                                                                                                                                                                                                          (BASIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NOTC_DROME STANDARD; PRT; 2703 AA. P07207; P04154; 01-NOY-1986 (Rel. 03, Created) 01-FB-1996 (Rel. 33, Last sequence update) 01-CCT-2000 (Rel. 40, Last annotation update) NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.
                 YEAST AND MAMMALIAN REGULATORY PROTEINS.
                                                                                                                                                                                                                                                       ARG/LYS-RICH (GLN-RICH.
ASN-RICH.
GLY/ASN-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1147 MOMNPNAAGPTAQOMMMGSGAGGPGQVPGP 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         535 SSVN---AAPT-----PGP 545
                                                                                                                                                                                                                                                                                                                                                                              GLN-RICH.
5 X 2 AA
                                                                                                                                                                                                                                                                                                                    GLN-RICH.
                                                                                                                                                                                                                                                                                                                                    GLY-RICH
                                                                                                                                                                                                                                                                                                                                                  GLN-RICH
                                                                                                                                                                                                                                                                                                                                                                 GLN-RICH
                                                                                                                                                                                                                           protein;
                                                                                                                                                               EMBL; X54251; CAA38152.1; -. PIR; A33106; A33106. PIR; A36391; A36391.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     167717
                                                                                                                                                                                                           FlyBase; FBgn0002643; mam.
Neurogenesis; Nuclear prot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             68; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                  407
651
700
759
987
1060
11092
11237
1492
1559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
Matches 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
NOTC_DROME
                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
```

```
EGF-LIKE 29.
EGF-LIKE 29.
EGF-LIKE 29.
CALCIUM-BINDING (POTENTIAL).
EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
EGF-LIKE 33.
EGF-LIKE 33.
EGF-LIKE 34.
EGF-LIKE 35.
                                            R InterPro; IPR000152; ...
R InterPro; IPR000152; ...
R InterPro; IPR0000561; ...
R InterPro; IPR001438; ...
R InterPro; IPR001488; ...
R InterPro; IPR001488; ...
R InterPro; IPR0011881; ...
R Pfam; PF00008; EGF; 36.
R Pfam; PF000010; EGFBLOOD.
R PRONITE; PS00010; EGFBLOOD.
R PROSITE; PS00010; EGFBLOOD.
R PROSITE; PS00010; EGFBLOOD.
R PROSITE; PS00010; EGFL; 34.
R PROSITE; PS00108; EGFL; 34.
R PROSITE; PS01186; EGFL; 34.
R PROSITE; PS01186; EGFL; 34.
R PROSITE; PS01186; EGFCA; 22.
R PROSITE; PS01187; EGFCA; 23.
R PROSITE; PS01187; EGFCA; 24.
R PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGF-LIKE 2.

EGF-LIKE 3.

EGF-LIKE 3.

EGF-LIKE 4.

EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 6.

EGF-LIKE 1.

EGF-LIKE 1.

EGF-LIKE 11.

EGF-LIKE 11.

EGF-LIKE 12.

EGF-LIKE 13.

EGF-LIKE 13.

EGF-LIKE 14.

EGF-LIKE 15.

EGF-LIKE 16.

EGF-LIKE 17.

EGF-LIKE 17.

EGF-LIKE 18.

EGF-LIKE 19.

EGF-LIKE 20.

EGF-LIKE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NEUROGENIC LOCUS NOTCH PROTEIN.
EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CYTOPLASMIC (POTENTIAL).
EGF-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-GLN (OPA-REPEAT).
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIN/NOTCH 1.
LIN/NOTCH 2.
LIN/NOTCH 3.
LIN/NOTCH 3.
ANK 1.
ANK 3.
ANK 4.
ANK 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glycoprotein
  HSSP; P00740; 11XA.
FlyBase; FBgn0004647; N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1052
1060
1098
1136
1183
1221
1259
1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1337
1375
1415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           514
554
901
950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ransmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Harris W.A.; "Many cell types specified by Notch function."; "Many cell types specified by Notch function."; Curr. Biol. 1:120-122(1991).
                                                                                                                                                                                                                                                                                                                                                                                               MEDIINE-87064624; PubMed-3097517; ... Kidd S., Kelley M.R., Young M.W.; "Sequence of the norch locus of Drosophila melanogaster: relationship of the encoded protein to mammalian clotting and growth factors."; Mol. Cell. Biol. 6:3094-3108(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 2505-2611 FROM N.A.
MEDLINE-85099329; PubMed-2981631;
Wharton K.A., Yedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;
"opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in D. melanogaster.";
Cell 40:55-62(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-8 FROM N.A.
MEDLINE-07257846; Pubmed-31037327;
KRelley M.R., Kidd S., Berg R.L., Young M.W.;
"Restriction of P-element insertions at the Notch locus of Drosophila
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Wooptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBL_TaxID=7227;
                                                                                                                                           SEQUENCE FROM N.A.
MEDILINE-866079539; PubMed-3935325;
Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
Whortion K.A., Johansen E.M. the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ECTODERM.
--- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
--- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -:- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
-:- SIMILARITY: CONTAINS 36 EGF-LITE DOMAINS.
-:- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
-:- SIMILARITY: CONTAINS 6 ANK REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M16152; AAB59220.1; -.
EMBL; M16149; AAB59220.1; JOINED.
EMBL; M16149; AAB59220.1; JOINED.
EMBL; M16151; AAB59220.1; JOINED.
EMBL; K03508; AAA28725.1; JOINED.
EMBL; K03508; AAA28725.1; JOINED.
EMBL; K03507; AAA28725.1; JOINED.
EMBL; M12175; AAA74496.1; -.
EMBL; M16025; AAA28726.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                melanogaster.";
Mol. Cell. Biol. 7:1545-1548(1987).
                                                                                                                                                                                                                                                                                                    Cell 43:567-581(1985)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A24420.
A24768.
A05267.
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. STRAIN-OREGON-R;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A24420; B
A24768; B
A05267; B
                                                                                                                                                                                                                                                                             repeats
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REVIEW.
```

GGANGGGVVGGGGGGGGGGGGPQPNSPVSLGIISPTGSDMGIMLAPPQSSKNSALMQTIS 2536

----KNPENQKVN---IDANGN-----AIPSGELXDDIVEQIA 322

292 2477 323 2537

383 TIALHRRNQPAEQT------TTTTTH-----TVVQQQTGG-IPQHKVALM 420 SEHSGOMNPPSIQSSMSGSSPSTNMLSPSSQHNQOAFYQYLTPSSQHSGGHTPQHTVQTL

2590

421

QQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGV 382

PQERRRESDRRDSQGSVASTHWSDSSSEVVNPYAE--VGGARNSL-----SAHQPEEHI 472

```
δ
                                                              ρp
                                                                                   QY
                                                                                                                οg
                                                                                                                                                δ
                                                                                                                                                                       염
                                                                                                                                                                                                   ÓΥ
                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 2703;
                                                                                                                                                                                                                                                                                                          SIMILARITY. SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY. SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY. SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY.
SIMILARITY.
SIMILARITY.
SINILARITY.
SINILARITY.
SIMILARITY.
                                                                                                                                         SIMILARITY. SIMILARITY.
                                                                                                                                                                                                            LARITY.
                                                                                                                                                                                                                                                                                                                                                  SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                 LARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 11;
                                                                                                               SIMILARITY
                                                                                                                                                                                  LARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 130;
                                                                                                                                                                                                                SIMI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.6%;
Query Match
Best Local Simil
Matches 120; C
 DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCULFID
DISCUL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
```

Search completed: September 27, 2001, 14:22:50 Job time: 146 sec

28;

QELLNGQGLGMNGNGQRNGVGPGVLPGGLCGMGGLSGAGNGNSHEQGLSPPYSNQSPPHS 2423

SD------PKFWVSVGAIAAGLAGL----AATG-----PKFWVSVGAIAAGLAGL

----EPDDPTTTDPDQAANAAESATKDQLTQEAF---- 291

IAQALALTP---

253

: | | | : | | | : | | : | : | : | : | SNAMAAPLANGNP-NTGAKQPPSYEDCIKNAQSMQSLQGNGL-DMIKLDNYAYSMGSPFQ 2363

AEARTRILAKLDPDNHGGRQPKDVD-----TRSVGVGSASGIDDGVVSETHTSTTNSSVR

AAAAVAAMSHELEGSPVGV-------GMGGNLPSPYDTSSMY 2305

ADGTHAAIG-EKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDI 174

: | | : :: || || :| || 1 || :| || || || TSAASKKAANLNGL--NDGQLT------GGVSGVPGVPPTN------SAAQAAAA 2270

2230

qq ò

116

26

õ

ŏ

2271

g &

2306 230 2364

g ŏ qq

175

SMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQ 115

GNLGNNVNGNHLIPPAPPLPSQT---DGAAR-----GGTGHLISSTGALGSRSLFSPLRN 55

81; Mismatches 194; Indels

Conservative

Similarity

THIS PAGE BLANK WERD,

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

September 27, 2001, 14:21:49; Search time 38.06 Seconds (without alignments) 1908.446 Million cell updates/sec Run on:

US-09-189-415A-2 2800 1 MPIGNLGNNVNGNHLIPPAP.....GETAVSSVNAAPTFGPVRFV 549 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

425026 seqs, 132305027 residues Searched:

425026 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SPTREMBL_16:* Database :

sp_invertebrate:*
sp_mammal:*
sp_mhc:* 1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:* sp_organelle:* sp_phage:* sp_plant:* sp_rodent:* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

sp_unclassified:*
sp_vertebrate:*
sp_virus:*

SIMMARIES

လ	Description	O50190 escherichia	052147 escherichia	Q9kwh9 escherichia	O68258 escherichia	Q9wxk1 escherichia	Q9etil citrobacter	Q47016 escherichia	Q47014 escherichia	085508 escherichia	Q9r396 escherichia	085506 escherichia	097169 drosophila	Q9n1p0 bos taurus	Q9fda0 xanthomonas	Q9nhx4 drosophila	Q9n5k0 caenorhabdi	Q9kxa6 escherichia	Q9t1k9 bacteriopha	Q9vn59 drosophila
SUMMARIES	Ω	050190	052147	Q9KWH9	068258	Q9WXK1	Q9ETI1	047016	047014	085508	Q9R396	085506	097169	Q9N1P0	Q9FDA0	Q9NHX4	Q9N5K0	Q9KXA6	Q9T1K9	09VN59
	DB	7	7	7	~	7	7	~	7	7	~	7	Ŋ	ဖ	7	'n	S	~	σ	Ŋ
	% Query Match Length DB	549	550	552	551	547	547	538	538	538	558	558	1158	818	1265	1323	1203	2806	2806	934
	% Query Match	99.8	92.6	89.3	78.8	74.2	74.1	63.6	63.6	63.0	54.6	54.2	6.0	5.7	5.7	5.6	5.4	5.4	5.4	5.4
	Score	2795	2676.5	2496.5	2206	2077	2076	1781.5	1781.5	1763.5	1528.5	1516.5	168.5	160.5	159.5	156.5	152	151.5	151.5	151
	Result No.	1	7	e	4	S	9	7	æ	9	10	11	12	13	14	15	16	17	18	19

084462 chli Q9xjml bac 054356 mor:		Q22248 caenorhabdi Q9v889 drosophila	Q9vtk8 drosophila O96506 drosophila	_	Qyneos leisnmania O14273 schizosacch	Q9n993 leishmania	Q9v358 drosophila	_	Q9u129 leishmania		Q9vp05 drosophila	Q9vxz1 drosophila	Q9erc7 mus musculu	Q9gyw7 drosophila	Q9i2m3 pseudomonas	Q9p319 neurospora	086808 streptomyce	Q9g0h8 roseophage	
084462 Q9XJM1 054356	034330 P91365 Q9GYB2	Q22248 Q9V889	Q9VTK8 Q96506	062672	Q9NE65 014273	09N993	Q9V358	Q9HAN2	Q9U129	P71401	Q9VP05	Q9VXZ1	Q9ERC7	Q9GYW7	Q912M3	Q9P3I9	.086808	0960н8	Q9KQ73
900	4 W W	വവ	വവ	ωı	ი ო	വ	ស	4	ഗ	~	Ŋ	Ŋ	Ξ	2	Ċ	'n	N	σ	7
1005 2806 832	2232 705	1829 848	2586 1296	1589	647	929	1472	1064	1860	2353	1455	2857	1066	1475	2468	770	1829	540	674
70.00 r.	ກິດ	5.2	2.5	•	, 5.1	2.0	2.0	5.0	5.0	4.9	4.9	4.9	4.9	6.4	4.9	4.9	4.9	4.9	4 .9
150 149.5		146 145.5	145.5	144	143.5	140.5	140.5	139	139	138.5	138	138	137	137	137	136.5		136	136
5 5 5 7 7 7 7	2 2 2 2 4 2 4 2	25 26	27	29	31 31	32	33	34	32	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

ö

a

```
181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 15, (TrEMBLrel. 15, (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               501; Conservative
                                                                                                                                                                                                                                      537 VNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PR003536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              552 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                         Abe A., Nagano H.;
                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                     Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PR01370;
                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                  01-OCT-2000
01-OCT-2000
01-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-HK01;
                                                                                                                                                                                                                                                                                                                                                                                          Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
                                                                                                                                                                                                                                                                                                о9кмн9
                                                                                                                                                                                                                                                                                                           Q9KWH9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                 418
                                                               301
                                                                                                       361
                                                                                                                                                                                                                                                                                      09KWH9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                        g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                               q
                                                                                                                                                                Dp
                                                                                                                                                                                     ò
                         ŏ
                                         g
                                                               ò
                                                                                 q
                                                                                                       ò
                                                                                                                          g
                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
Lai L.C., McNamara B.P., Donnenberg M.S., Kaper J.B.;
MAO1. Microbiol. 0.0-0(1998).
EMBL; AF022236; AAC38390.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVGADGTH 120
                                                                                                                                               420
                                                                                                                                                                                                                            GYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNAA 540
                                                                                                        360
                                                                                                                                                          540
                                                                           PQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAADP 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                                                                                                        SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALM
                          181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA
                                                                 241 IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN
                                                                                                         IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL
                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 2676.5; DB 2; Length 550;
Pred. No. 5e-148;
4; Mismatches 11; Indels 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19DD08A9BE9251CB CRC64;
                                                                                                                                                                                                                                                                                                                                                            01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                            Ą
                                                                                                                                                                                                                                                                                                                                           550
                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1370; TRNSINTIMINR.
550 AA; 56509 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 95.6
Best Local Similarity 96.0
Matches 531; Conservative
                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR003536;
                                                                                                                                                                                                                                                                                  |||||||||||||||
PTPGPVRFV 549
                                                                                                                                                                                                                                                                       PTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                 Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBL_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PR0137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS; PI
SEQUENCE
                                                                                                                                                                                                                                                                                                                                            052147
                                                                                                                                                                                                                                                                                                                                                       052147
                                                                                                                                                                                                                                                                                                                        7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                     361
                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                        541
                                                                                                          301
                                                                                                                              301
                                                                                                                                                                                                                                 481
                                                                                                                                                  361
                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                             g
                                                                                                                                                                 g
                                                                                                                                                                                         ò
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                        ò
                           ò
                                             g
                                                                   ò
                                                                                       g
                                                                                                         ò
                                                                                                                                                  ò
```

```
AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Analyses of type III secreted proteins and Tir in enteropathogenic Escherichia coli 0157:H45.";
Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB036053; BAA96815.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS
                                                                                                                                                                             AADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSS
                                                                                                                                                                                                                                                                                                                                  241 IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN
                                                                             IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL
                                                                                                                                                           SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALM
                                                                                                                                                                                                                                                         GTDATRAEDASLNRRDSGGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEV
                                                                                                                                                                                                                                       421 PQERRRFSD----RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 552;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRNSINTIMINR.
A: 57005 MW; A8D79EE22EE50A4B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89.2%; Score 2496.5; DB 2
90.3%; Pred. No. 1.4e-137;
tive 11; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
```

٠. ښ

m

g á q δ 셤 δ g ò g à 셤

```
SSGIGYGLSSALIVAGGIGAGVTTALHRRNOPAEQTTTTTTTTTVVQQQTGGIPQHKVALM 420
                                                                                               361 SSGIGYSLSSALIVGGGIGAGVTTALHRRNQPAEQTTTTTHTVVQQQTGG---NTPAQG 417
                                                                                                                                         PQERRRESD----RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEV 476
                                                                                                                                                         477 AADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLA-NSGGLRLGMGGLTSGGETAVS 535
                                                                                                                                                                                                               61 VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241 IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN 300
241 IAAGLAGLAATGITQALALTPEPDDPTTTDPEQAASAAESATRDQLTQEAFKNPENQKVS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=MPEC;
Okutani A., Itoh K., Sasakawa C.;
"Translocated intimin receptor(Tir) of murine pathogenic Escherichia
                                         301 IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    547;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      coli 0115a,C:K(B).";
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB026719; BAA77400.1; -.
InterPro; IPR000484; -.
InterPro; IPR003536; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       02CAC6D625FA6EE1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-NAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  74.2%; Score 2077; DB 2; 76.0%; Pred. No. 3.5e-113;
                                                                                                                                                                                                                                                                                                                                                           547 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44; Mismatches
                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PR01370; TRNSINTIMINR.
ProDom; PD000551; -; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          56283 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                          536 SVNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                         : |||||||| ||| 538 TANAAPTPGPGREV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          547 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 418;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Receptor.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                            09WXK1
                                                       301
                                                                                                                                         421
                                                                                                                                                                                                                                                                                                                                RESULT
Q9WXK1
                                                                                                        g
                                                                                                                                                                   g
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                       δ
                                                    a. y
                                                                                                                                         ŏ
                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN 300
 IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN 300
                                                       IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL 360
                                                                                                              SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTVVQQQTGGIPQHKVALM 420
                                                                                                                             PQERRRFSD----RRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEV 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Voss E., Paton A.W., Manning P.A., Paton J.C.;
"Molecular analysis of Shiga toxigenic Escherichia coli 0111:H-
proteins which react with sera from patients with hemolytic-uremic
syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6EC95F76EF0F44CC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78.8%; Score 2206; DB 2;
79.8%; Pred. No. 1.1e-120;
tive 39; Mismatches 65;
                                                                                                                                                                                                                                                                                                                                                                                          551 AA
                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE-98187918; Pubmed-9529069;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Infect. Immun. 66:1467-1472(1998)
EMBL; AF025311; AAC69249:1; -
InterPro; IPR003536; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PR01370; TRNSINTIMINR.
E 551 AA; 56975 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                     549
                                                                                                                                                                                                                                                                                                                 552
                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                     SSVNAAPTPGPVRFV
                                                                                                                                                                                                                                                                                                    538 ANANAAPTPGPVRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 442; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Escherichia coli,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                       068258
                                                       301
                                                                                                                                                                                                                                                        478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                 301
                                                                                                            361
                                                                                                                                         361
                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                     535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                              477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
```

ä

Gaps

4 ;

ò 셤 ö g ò g ŏ g ò

9

240

238

```
480 PGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNA 539
                                            IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL 360
                                                            361 SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTHTVVQQQTGGIPQHKVALM 420
                                                                                                                     421 PQERRRES-DRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases. EMBL; 159504; AAD19750.1; -. EMBL; AF045568; AAC15683.1; -. EMBL; U59503; AAB02941.1; -. Interpro, IPR003536; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=RDEC-1;
Agin T.S., Boedeker E.C.;
Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TIEMBLEAL: 01, Created)
01-AUG-1999 (TIEMBLEAL: 11, Last sequence update)
01-AUR-2001 (TIEMBLEAL: 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                            Æ
                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PR01370; TRNSINTIMINR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 172-538 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         to the
                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 65.4
Matches 367; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       O'Brien R.A.;
Submitted (DEC-1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Krejany E.O.;
Submitted (MAR-1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-REPEC RDEC-1;
                                                                                                                                                                                                                                                                                                                   547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. STRAIN-REPEC 83/39;
                                                                                                                                                                                                                                                                                        540 APTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. STRAIN-REPEC 83/39,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Escherichia coli
                                                                                                                                                                                                                                                                                                                   538 ATTPGVERFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Receptor.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                              047016
                                                                                                                                                                                                                                                                                                                                                                                                               047016
                                                                                                                                           359
                                                      301
                                                                                                                                                                                                                                                                                                                                                                  RESULT
Q47016
                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                             δ
                                                                                                                                         g
                                                                                                                                                                        ò
                                                                                                                                                                                               qq
                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                   g
                                                    δ
                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-CDC 1843-73T, AND DBS100;
MEDLINE-20553330; PubMed-11101562;
MEDLINE-20553330; PubMed-11101562;
Luperchio S.A., Newman J.V., Dangler C.A., Schrenzel M.D.,
Luperchio S.A., Newman J.V., Schauer D.B.;
Brenner D.J., Steigerwalt A.G., Schauer D.B.;
Clirobacter rodentium, the Causative Agent of Transmissible Murine
Colonic Hyperplasia, Exhibits Clonality: Synonymy of C. rodentium and
Mouse-Pathogenic Escherichia coll.";
Clin. Microbiol. 38.4343-4350(2000).
EMBL; AF301618; AAG40758.1;
EMBL; AF301617; AAG25642.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 ILAKLDPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVSVGA 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVN 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 VDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTH 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         361 SSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALM 420
                                                                                                                                421 PQERRRES-DRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHIYDEVAAD 479
                                                                                                                                                                                           PGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSGGETAVSSVNA 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                         301 IDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEELQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2; Length 547;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74.1%; Score 2076; DB 2; Length 5
76.0%; Pred. No. 4e-113;
Live 44; Mismatches 84; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          547 AA; 56270 MW; CB8318B301049C37 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
11-MAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                     547 AA
                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 418; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Citrobacter rodentium
                                                                                                                                                                                                                                                                                                  540 APTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                              538 ATTPGVERFV 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=67825;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Citrobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Receptor.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                     Q9ETI1
Q9ETI1;
```

9

8

ð g φ

ò g 엄 ò

ŏ

```
1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                           Length 538;
                                                        63.6%; Score 1781.5; DB 2; Length
65.4%; Pred. No. 5.1e-96;
ive 66; Mismatches 93; Indels
538 AA; 55410 MW; 3132A969B7B3D06C CRC64;
```

241

g δ g ò Оp ογ 셤 δ

S

us-09-189-415a-2.rspt

```
FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  085508;
085508;
01-NOV-1998 (
01-NOV-1998 (
                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
085508
ID 08
AC 08
DT 01
DT 01
RET REP REP RED DR RET 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                         ||: ||:| | ::| | ::| |:| | ::| || ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 VDSR--DIPGLPINPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG 118
                                                                            60 ADSRASDIPGLPTNPLRFAA--SEVSLHGALEVLHDKGGLDTLNSAIGSSLFRVETRDDG 117
                                                                                                                                                                                              THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR 178
                                                                                                                                                                                                                                        179 TRILAKLDP-DNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVS 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQ 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEE 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTTTTVVQQQTGGIPQHKV 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             418 ALM-----PQERRRFSDRRDSQGSVASTHWSDSSS--EVVNPYAEVGGARNSLSAHQP 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SAQGNTDTSGPEESP--ASRRNSNASLAS-NGSDTSSTGTVENPYADVGMPRNDSLARIS 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDILINE-59903184; PubMed-9784578;
Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
"Translocated intimin receptors (Tir) of Shiga-toxigenic Escherichia
"Translocated intimin receptors (Tir) of Shiga-toxigenic Escherichia
"Translocated intimin receptors (Tir) of Shiga-toxigenic Escherichia
"Translocated intimin receptors" (Tir) of Shiga-toxigenic Escherichia
"Translocated intimi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-413/89-1 (026:H-);
MEDLINE-98294040; PubMed=9632251;
Deibel C., Kraemer S., Chakraborty T., Ebel F.;
"Eppp, a novel secreted protein of attaching and effacing bacteria, directly translocated into infected host cells, where it appears as tyrosine-phosphorylated 90 kDa protein.";
Mol. Microbiol. 28:463-474(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1996 (TrEMBLrel. 01, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN-REPEC 84/110/1, AND E65/56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGETAVSSVNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Krejany E.O.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    047014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             047014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          469
                                                                                                                                                                                              119
                                                                                                                                                                                                                                                                          118
                                                                                                                                                                                                                                                                                                                                                                                                                                       178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTTHTVVQQQTGGIPQHKV 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              517
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                          59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                                              61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179 TRILAKLDP-DNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALM-----PQERRRFSDRRDSQGSVASTHWSDSSS--EVVNPYAEVGGARNSLSAHQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTS
                                                                                   :
                                                                                                                                                                                                                                                                                                                                                                                           35;
                                                               Boullier S., nougayrede J.-P., Boury M., Mainil J., Charlier Boullier S., De Rycke J., Milon A., Oswald E.; "Role of Tir and Intimin in the pathogenesis of rabbit enteropathogenic Escherichia coli."; Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases. EMBL, J05902; AAG12028.2; -
                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                         31D7A8E227B3D06C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                     63.6%; Score 1781.5; DB 2;
65.4%; Pred. No. 5.1e-96;
iive 66; Mismatches 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
sequence heterogeneity.";
Infect. Immun. 66:5580-5586(1998).
                                                                                                                                                                                    AJ223063; CAA11065.1; -.
AF070068; AAC69316.1; -.
AF132728; AAD27868.1; -.
AF113597; AAF03080.1; -.
                                                                                                                                                                                                                                                                      PRINTS; PR01370; TRNSINTIMINR. Hypothetical protein; Receptor. SEQUENCE 538 AA; 55420 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08,
08,
16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               529 GGETAVSSVNAAPTPGPVRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 518 GGESAVSTANAAPTPGPARFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best_Local Similarity 65.49
Matches 367; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. (TrEMBLrel. (TrEMBLrel. )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                    InterPro; IPR003536;
```

```
TRANSLOCATED INTIMIN RECEPTOR TIR
                    Escherichia coli
                                                                                                                                         phosphorylated.
                                                NCBI_TaxID=562;
                                                                                                                                                                                STRAIN-EDL933
                                                                               STRAIN-86/24;
                                                                                                                                                                                                                                                                                                                                         Best Local Sin
Matches 327;
                                                                                                                                                                                                                                                                                           Receptor.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
   g
                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTHTVVQQQTGGIPQHKV 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     418 ALM-----PQERRRFSDRRDSQGSVASTHWSDSSS--EVVNPYAEVGGARNSLSAHQP 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SAQGNTDTSGPEESP--ASRRNSNASLAS-NGSDTSSTGTVENPYADVGMPRNDSLARIP 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTS 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                     KVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQHARRQEE 357
                                                                                                                                                                                                                                                                                                                                 119 THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR 178
                                                                                                                                                                                                                                                                                                                                                                      TRILAKLDP-DNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTSTTNSSVRSDPKFWVS 237
                                                                                                                                                                                                                                                                                                                                                                                   VGAIAAGLAGLATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQ 297
                                                                                                                                                                                                                                                                                           61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG 118
                                                                                                                                                                                                                                                                                                       0157
                                                                                                                                                                                                                                                               1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS 60
                                                                                                                                                                                                                                     Gaps
                                                                                                              and
                                                                                       Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
"The translocated intimin receptors (Tir) of Shiga toxigenic
Escherichia coli isolates belonging to serogroups 026, 0111, and
react with sera from patients with hemolytic-uremic syndrome and
exhibit marked sequence heterogeneity.";
Submitted (Jun.1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF070069; AAC69318.1;
InterPro; IPR033536;
                           Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                              Query Match 63.0%; Score 1763.5; DB 2; Length 538; Best Local Similarity 64.5%; Pred. No. 5.7e-95; Matches 362; Conservative 71; Mismatches 93; Indels 35;
                                                                                                                                                                                  55602 MW; 447052A0E3214D6D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      558 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
       TRANSLOCATED INTIMIN RECEPTOR TIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   529 GGETAVSSVNAAPTPGPVRFV 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS; PR01370; TRNSINTIMINR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 13, TremBLrel. 13, TremBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                       538 AA;
                                                                            SEQUENCE FROM N.A.
                                                        NCBI_TaxID=562;
                                                                                    STRAIN-EPEC87A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY-2000
01-MAY-2000
01-MAR-2001
                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
Q9R396
ID Q9R396
AC Q9R396
DT 01-MAY
DT 01-MAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09R396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9R396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                358
                                                                                                                                                                                                                                                                                                                                                                             179
                                                                                                                                                                                                                                                                                                                                                                                                                                      227
                                                                                                                                                                                                                                                                                                                                                                                                                                                          298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                            da
Oy
                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                     셤
         g
                                                                                                                                                                                                                                                                                                 à
```

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FKNPENOKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQ 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   351 HARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTT-----HT 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSS-SEVVNPYAEVGGARNS 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GRLIGTPGQGIQ 507
                                                                                                                                                                                                                                                              290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      179 TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGI-----DDGV--VSETHTSTTNSSVRS 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     177 ORILELLEPKGTG------ESKGAGESKGVGELRESNSGAENTIETQTSTSTSSLRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     įs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS
                                                                                                                                   Finlay B.B.;
"Enterohemorrhagic Escherichia coli 0157:H7 produces Tir, which
translocated to the host cell membrane but is not tyrosine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      57;
Bacterla; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2; Length 558;
                                                                                           MEDLINE-99242825; Pubmed-10225900;
DeVinney R., Stein M., Reinscheid D., Abe A., Ruschkowski S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54.6%; Score 1528.5; DB 2; Length
56.2%; Pred. No. 2.7e-81;
tive 60; Mismatches 138; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                          558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVT----
                                                                                                                                                                                         Infect. Immun. 67:2389-2398(1999)
                                                                                                                                                                                                                                                 эткати~вынэээ;
MEDLINE=98339885; PubMed=9673266;
                                                                                                                                                                                                                                                                                                                                                                                                 PRINTS; PR01370; TRNSINTIMINR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                      SEQUENCE FROM N.A.
```

ò g

```
Query Match
Best Local Similarity 18.99
Matches 130; Conservative
                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                               FROM N.A.
                                                                                                                                                                                                                                                                   HAIRLESS PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             I-----
                                                                                                                                                                         097169;
01-MAY-1999 (
01-MAY-1999 (
01-MAR-2001 (
                                                                                                            12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              305
                                                                                                            RESULT
                                                                                                                               097169
ID 09
                                     g
                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VDSR--DIPGLPINPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADG 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              THAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEAR 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSS-SEVVNPYAEVGGARNS 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    177 ORILELLEPKGTG-----ESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQAQQRYEDQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTT-----HT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVT------GRLIGTPGQGIQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRILAKLDPDNHGGRQPKDVDTRSVGVGSASGI-----DDGV--VSETHTSTTNSSVRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
"The translocated intimin receptors (Tir) of Shiga toxigenic
"The translocated intimin receptors (Tir) of Shiga toxigenic
scherichia coli isolates belonging to serogroups 026, 0111, and
react with sera from patients with hemolytic-uremic syndrome and
exhibit marked sequence heterogeneity.";
Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF070067; AAC69314.1; -.
                                                                                                                                                                                                                                                                                                                        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2; Length 558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54.2%; Score 1516.5; DB 2; Length 55.8%; Pred. No. 1.3e-80; Live 59; Mismatches 141; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CA2CDDAC94527C2E CRC64;
                       Last sequence update)
Last annotation update)
STYALLANSGGLRLGMGGLTSGGETAVSSVNAAPTPGPVRFV
                                                                                                                                                        Æ
                                                                                                                                                                                                     Created)
                                                                                                                                                      PRT;
                                                                                                                                                                                           01-NOV-1998 (TrEMBLrel. 08, Creates 01-NOV-1998 (TrEMBLrel. 08, Last ss 01-MAR-2001 (TrEMBLrel. 16, Last st TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PR01370; TRNSINTIMINR. SEQUENCE 558 AA; 58175 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 55.8
Matches 325; Conservative
                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-562;
                                                                                                                                                                                                                                                                                                               Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                        RESULT 11
085506
AC 085506;
DT 01-NOV-
DT 01-NOV-
DT 01-MAR-
DE TRANSLO
GN TIR.
OS BSCHETI
OC BACTETI
OC BACTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      408
```

ò g ŏ q ö ద á g ö g ò g Óλ qq οy g ò g

```
Drosophila hydei (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBL_TaxID=7224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :: | | ; : : | | ; : : | | ; : : | | ; : : | | ; : : | | ; : : | | ; : : | | ; : : | | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : : | ; : | ; : : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ; : | ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SRSLFSPLRNSMADSVDSRDIPGLPTNPSRLAAATSETCLL-GGFEVLHDKGPLDILNTQ 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---GHPMVTVASD-----RQPKDV 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVEQIAQQAKEAGEVARQQAVESNAQ 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AQQRYEDQHARRQEE-LQLSSGIGYGLSSALIVAGGIG------AGVTTA----LHR 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNQPAEQTTTTTT-----KVALMPQ 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TGTPATKTTTMMTDIRLKKOKPRAKLNSIIOKLIDGVPARLEQLSKTPAVAATASAVAS 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::::|:| | | | | | |::
117 TSNGTTPALSSTKTNSSSNS----SSSSLIMATASATALVAGGAGVTAPKAPADVMAGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LKNGINGTNTAAAAAATLLTAAAAVTFQLPAAPELKPTTSATAANATNSTTSSSTATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 105 IGPSAFRVEVQADGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DMGRHPISMRSNNSMSGYGGRLQFFKDGKFILE-----LARAKDGDKGGWVSVPRKPFRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DTRSVGVGSASGIDDGVVSETHT -----STTNSSVRSDP----KFWVSVGAIAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LGNNVNGNHL--------IPPAPPLPSQTDGAARGGTGHLISSTGALG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :: | | : | : | : | : | : | : | | PSAAT---SATVTPTSAVTTYPKNENSTSLSFSDDNSSIQSSPWQRDQPWKQT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LAGLAATGIAQALAL------TPEPDDPT-TTDPDQAANAAESAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGARNSLSAHQPEEHI-----Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-WILD TYPE. CV. HESS, O.; Mater D.; Marquart J., Alexief-Damianof C., Preiss A., Maler D.; Marquart J., Alexief-Damianof C., Preiss A., Maler D.; Rapid divergence in the course of Drosophila evolution reveals structural important domains of the Notch antagonist Hairless."; Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases. EMBL, AJ23783, CAB38221.1; FlyBase: FBQUGD199; DhydVH. SEQUENCE 1158 AA; 119635 MW; OB611D3739CF397F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches 238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      υ
:
                                                              STYALLANSGGLRLGMGGLTSGGETAVSSVNAAPTPGPVRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 168.5; DB
Pred. No. 0.079;
                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (TrEMBLrel. 10, (TrEMBLrel. 10, 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.0%;
18.9%;
```

```
NCBI_TaxID=64187;
                                                                                                                                                                                                                                                                                                                      Xanthomonas.
                                                                                                                                                                                                          Q9FDA0;
                                                                                                                                                                                             Q9FDA0
                                       461
                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               446
                                                                                                                         589
           496
                                                                                                                                                                                Q9FDA0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                 g
                                                                                              ğ
                                                                                                                      g
           g
                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          323 PGSTQGG-----SAATGGSGAGSGPTAPVSGETRTSVISGINVPVSGAP----VTPGSSA 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           374 GSSGAPGTGGPGSETASPLSGAAGTSATGSGTSI---PPSGAPVTPEPPLLSTGA-SAGP 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              403 TVVQQQTGGIPQHKVALMP--QERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGGAR 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        70 PTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQADGTHAAIGEKNGL 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             130 EVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVASDIAEARTRI----LAKL 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186 DPDNHGGRQPKDVDTRSVGVGSASGIDDGVVSETHTST---TNSSVRSDPKFWVSVGAIA 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -------AGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATK 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        343 AQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTH 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  228 PGSTTGRAAGAGTPA------VDSQQTASLPAAARPTALGPGTSAPSGETSES 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         275 RSSVP-----GGSETTQQPG----AGSESPTLSPGVTRTTALRGSETRVPSTGVSGL 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  597
                                            ----GGISGGAGGKRSRAKGS 633
                                                                                                                                                                                                                                                                                                                                                                                                                     20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10 VNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSMADSVDSRDIPGL 69
                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
STRAINBERED ANGGS.
MEDLINE-20223253; PubMed-10759843;
MEDLINE-20223253; PubMed-10759843;
MEDLINE-2022353; PubMed-10759843;
Jiang W., Gupta D., Gallagher D., Davis S., Bhavanandan V.P.;
The central domain of bovine submaxillary mucin consists of over 5 tandem repeats of 329 amino acids: chromosomal localization of the BSM1 gene and relations to ovine and porcine counterparts.";
EMBL, J. Biochem. 267:2208-2217(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  284 DQLTQEAFKNPENQKVNIDANGNAIP-SGELXDDIVEQIAQQAKEAGEVARQQAVESNAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            430 PASSESTVTLPGATGTDVLRSGTSLPVSG-----GAVTPAPSPGGSSA-----
                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
540 TADRISSNSSGGGAIGSL -- SHSLAHKVSPPSSAAAASRLVEYHHHHQHVSPRKRILREF
                           DEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTS--GGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 818;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B5BB44F84F66F86B CRC64;
                                                                                                                                                                                                          01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
SUBMAXILARY MUCIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 5.7%; Score 160.5; DB 6; Best Local Similarity 21.6%; Pred. No. 0.15; Matches 122; Conservative 64; Mismatches 221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----TAGPGVGSATTVQASGATGADVL
                                                                                                                                                                                  818 AA
                                                       EKVSLEDNNGCVNNGSGVG------
                                                                                  TAVSSV-----NAAP----TPGP 545
                                                                                                              634 TATSAVTTKSMPINLAPPQAKVLSTPTP 661
                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        818
73014 MW;
                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                          Mammalia; Eutheria; Ce
Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                 taurus (Bovine).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        818
818 AA;
                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                Q9N1PO
                                                                                                                                                                                  0911PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    473
                                                        598
                            474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                       8
                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                              ð
                                                                                  à
```

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ray S.K., Rajeshwari R., Sonti R.V.;
"A putative outer membrane protein from Xanthomonas oryzae pv. oryzae that is involved in virulence.";
Submitted (JUL-2000) to the BMBL/GenBank/DDBJ databases.
EMBL, AF288222; AAG01335.1; ...
SEQUENCE 1265 AA; 119856 MW; 21762579B5EC70A2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GEVARQQAVESNAQAQQRYEDQHARRQEELQLSSGIGYGLSSALIVAGGIGAGVTTALH- 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSEV----VNPYAEVGGARNSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGT 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56 SMADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQ 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        116 ADGTHA-AIGE--KNGLEVSVTLSPQEWSSLQSIDTEGKNRFVFTGGRGGSGHPMVTVAS 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 696 ATGVSAVAIGEISKATGEESVAVGGGAFSGWIPTQASGKGAAAFGAGAWATADYTTAIGR 755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     816 IGENAVSVGYNSYVRQS----AVNGVALG-ANAGATG-ADSVAL---GSGSSTYDADTV 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  275 ANAAESATKDQLTQEAFKNPENQKVNI-----DANGNAIPSGELXDDIVEQIAQQAKEA 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        547
                                                                         520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 IGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISS-TGALGSRSLFSPLR-----N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --RRNQPAEQTTTTTTTTTTTVVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            935 ASYNNVGAALTALDSKVTELDARSGGTPANTAA-----RTASLRTATVPAMAATAVSAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           866 SVGSGNGRGGPATRI-----VNVGAGAVASASTDAINGGQLFESL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1/3 D-IAEA--RTRILAKLD--PDN----HGGRQPKDVDTRSVGV-GSASGIDD-----GV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           216 VSETHTST-TNSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQA
NSLSAHQPEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIGTPGQGIQSTYALLANSGGLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Xanthomonas oryzae pv. oryzae.
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match

5.7%; Score 159.5; DB 2; Length 1265;
Best Local Similarity 23.4%; Pred. No. 0.3;
Matches 136; Conservative 73; Mismatches 229; Indels 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEWBLrel. 16, Created)
01-MAR-2001 (TrEWBLrel. 16, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
PUTATIVE OUTER MEMBRANE PROTEIN XADA.
                                                                                                                                                                                                                                                                                                                                                                                                                          Ą.
                                                                                                                                                                                                                                             521 LGMGGLTSGGETAVSSVNAAPTPG
                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
```

g	989 SSNVASTAIDATAGVQGTPTAAVVGSITPAAISTVVGTAAVANNVTGTAIG- 1039	QY
ζ	502 PGGGIQSTYALLANSGGLRLGMGGLTS-GGETAVSSV 537	qa
ą	1040GSAYAHGANDIAIGSNARVNADGSTAVGANTQIAAV 1075	Qy
RESULT		qa
29NHX4	6	Qy
	(TrEMBLine) 15. Created)	qq
555	01-OCT-2000 (TrEMBLrel. 15, Last Sequence update)	č
SNE	ALHAMBRA OR CG1070. Drosophila melanogaster (Fruit fly).	Job t
8888	OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; OC Ephydroidea; Drosophilidae; Drosophila.	·
5 Z 2	SEQUENCE FROM N.A.	
RT R	Perrin L., Dura J.M.; "Alhambra, a Drosophila homolog of mammalian AF-10 and AF-17."; Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.	
888	EMBL; AF217960; AAF72595.1; FlyBase; FBgn0037471; Alhambra. Interpro; IPR001965;	
SOR	SMART; SM00249; PHD; 1. SEQUENCE 1323 AA; 132883 MW; D53COC8AF392F9A6 CRC64;	
Que Bes Mat	Query Match 5.6%; Score 156.5; DB 5; Length 1323; Best Local Similarity 20.2%; Pred. No. 0.47; Matches 135; Conservative 85; Mismatches 270; Indels 177; Gaps 25;	
b Sy	3 IGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSP-LRNS 56	
ζο qα	57 MADSVDSRDIPGLPTNPSRLAAATSETCLLGGFEVLHDKGPLDILNTQIGPSAFRVEVQA 116	
λα	DGTHAAIGEKNGLEVSVTLSPQEWSSLQSIDTE-GKNRFVFTGGRGGSG-HPM	
Q O	434 SSTREKSSSKLSKNKDSNQVPSATSSLSTTSSINTQPSSSTSTATAGSGGTGTHVS 489	
č q	168 VTVASDIAEARTRILAKLDPDNHGGRQPKDVDTRSVGSGSASGIDDGV	
λ G G	216VSETHTSTTNSSVRSDPKFWVSVGAIAAGLAGLAATGIAQALALITPE 262 	
à é	263 PDDPTTTDPDQAANAAESATKDQLTQEAFKNPENQKVNIDANGNAIPSGELXDDIVE 319	
ìò	O1AO-OAKEAGEVAROOAVESNAOAOORXEDOHARROEELOLSSG1GXGLSSA	
; 음	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::	
oy Op	372 LIVAGGIGAGVTTALHRRNOPAEQTTTTTTHTVVQQQTG 410 :	
ý á	411 GIPOHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAE 455 1	
3	GIPTIESSNSSGIASAGSGSSSNISSIAAGSNIALISSSNUUASIASSSARSLIVS	

Qγ	456	456 VGGARNSLSP	GGARNSLSAHOPEEHIYDEVAADP	480
q	831	831 VPLSTANVPGINLPTSSTSSSTTSESHSASSRSGAQSQHQQQLSNALVGPPMGGSAGAF 890	ASSRSSGAQSQHQQQLSNALVGPPMGGSAGAF	890
δλ	481	481GYSVIQNFSG-SGPVTGRLIGTPGQGIQSTYALLANSGGLRLGMGGLTSG 529	YALLANSGERLGMGGLTSG	529
qq	891	891 HGGTTSAGSSSVIQQQSGKSSPALGTLVSGNSGGSIISASGFPLPSG 937	GNSGGSIISASGFPLPSG	937
QY	530	530 GETAVSS 536		
qa	938	938 NLTATTT 944 .		
Searc	ch con	Search completed: September 27, 2001, 14:27:03		

```
WPI; 1999-337712/28.
N-PSDB; AAX58859.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Key
Misc-difference 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09924576-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L2-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-MAY-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY06221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY06221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mutant protease (d
Mutant protease (K
Mutant protease (N
Mutant protease (A
Mutant protease (A
                                                                                                                            ; Search time 31.79 Seconds (without alignments) 1066.021 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Intimin C-terminal
Human PRO1342 (UNQ
Human PRO1342. Ho
Protein of the inv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EHEC E. coli trans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                    US-09-189-415A-4
2851
1 MPIGNLGHNPNVNNSIPPAP......SNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                **Singlify of the control of the con
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                         412676 seqs, 60623988 residues
                                                                                                                                September 27, 2001, 14:22:24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY06221
AAY06220
AAB20576
AAB30576
AAB87575
AAB66157
AAR10562
AAR10561
AAR10561
AAR10551
AAR10551
                                                                                        OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A_Geneseq_0601:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1959
1962
1962
1962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2851
1498
406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        148.5
148.5
141.5
141.5
141.5
141.5
                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                                                     ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
No.
```

70	ALIGNMENTS					
κý	AAB69137	22	1992	4.3	21.	45
M. catarrhalis str	AAB69133	22	1992	4.3	121.5	44
oraxella	AAW04505	17	1992	4.3	21.	43
	AAY67238	21	456	4.3	21.	42
Amino acid sequenc	AAY37175	20	496	4.3	22.	41
. Host cell factor p	AAR57141	12	2035	4.3	12	40
Haemophilus adhesi	AAR99392	17	1098	4.3	2	33
EPEC E. coli trans	AAY06213	20	30	4.3	12	38
	AAW30292	18	1601	4.3	٠	37
M. tuberculosis re	AAY38990	20	359	4.3	C	36
M. tuberculosis an	AAY39133	202	359	. 4 	N	35
Mycobacterium tube	AAW64336	1 6	6 G	4.4	124	0 K
Neisseria gonorrhe	AAY75096	21	2048	4.4	24.	32
High molecular wei	AAR41732	14	1529	4.4		31
MSP encoded by pUC	AAR14150	12	461	4.4	25.	30
Enterococcus faeca	AAY00143	20	558	4.5	27.	29
Human cytomegalovi	AAW27277	18	1048	4.5	12	28
Amino acid sequenc	AAY35195	20	550	4.6		27
sp45 pro	AAR14530	12	461	4.6	32.	56
. cat	AAB69136	22	2314	4.7	133	25
gondii	AAY29082	20	288	4.7	m	24
T. gondii immunoge	AAY29081	20	288	4.7	13	23
	AAY95556	21	3178	4.7		22
Caenorhabditis ele	AAY95559	21	2870	4.7	33.	21
Neisseria meningit	AAY75097	21	2514	4.7	13	70
	AAY 00142	20	1638	4.7	34.	19
	AAY00140	20	1638	4.7	34.	18
ccus fae	AAY00138	20	1638	4.7	34.	17
prote	AAY83025	21	1026	4.9	40.	16
ďΩ	AAR10940	12	1974	5.0	41.	15
Mutant protease (d	AAR10941	12	1968	5.0	41.	14
proteas	AAR10563	12	1962	2.0		13
Mutant protease (A	AAR10559	12	1962	5.0	41.	12

541 nsavntsnnppapgshrfv 559

```
The present sequence represents Tir, a novel translocated intimin receptor (formerly termed Hp90) from an enterohaemorrhagic scherichia coll (EHEC) strain. The sequence was deduced from an isolated tir polynucleotide (see AAX58859). Tir proteins are secreted by attaching and effacing pathogens such as EHEC and EPEC (see AAY06220) E. coli. The bacterial pathogens insert their own creceptors into mammalian cell surfaces, to which the pathogen then adheres to trigger additional host signaling events and actin nucleation. Diagnosis of disease caused by pathogenic E. coli can be performed by use of antibodies that bind to Tir to defect the protein or the use of antibodies that bind to Tir to defect the protein or the use of nucleic acid probes for detection of nucleic acids encombinant method for producing recombinant Tir, and a kit for the detection of Tir-producing E. coli are provided. A method of immunising a host in addition, Tir fusion proteins can be used in attenuated E. coli in addition, Tir fusion proteins can be used in attenuated E. coli in didition, Tir fusion proteins can be used in attenuated E. coli interfere with the binding of bacterial pathogens to their perpendents is also provided. Interfere with the binding of bacterial pathogens to their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RINGIQISVVXSTIQHPPRDTTDNGARLLGNPSAGIQSTVARLALSGGLRHDMGGLTGGS 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        lgnaipsgylkddyvanieeqakaageeakooaiennaqaokkydeooakroeelkyssg 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPANNTPAQG 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLF 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVAT
         New translocated intimin receptor useful for treating infection by enteropathogenic or enterohemorrhagic Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Score 2851; DB 20;
; Pred. No. 9.3e-213;
0; Mismatches 0;
                                                               Claim 7; Page 55-58; 91pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                          receptors is further provided
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 100.
Best Local Similarity 100.
Matches 559; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           559 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
δ
```

```
The present sequence represents Tir, a novel translocated intimin receptor (formerly termed Hp90) from an enteropathogenic receptor (formerly termed Hp90) from an enteropathogenic scheduling to the sequence was deduced from an isolated tir polynucleotide (see AAX58858). Tir proteins are secreted by attaching and effacing pathogens such as EPEC and EHEC secreted by attaching and effacing pathogens such as EPEC and EHEC (see AAX06221) E. coli. The bacterial pathogens insert their own cacheres to trigger additional host signaling events and actin uncleation. Diagnosis of disease caused by pathogenic E. coli can nucleation. Diagnosis of antibodies that bind to Tir to defect the protein or the use of antibodies that bind to Tir to defect the protein or the use of nucleic acid probes for detection of nucleic acids encodaing Tir polypeptide. Isolated Tir nucleic acids, Tir peptides, a recombinant method for producing recombinant Tir, and a kit for the detection of Tir peptides, a recombinant method for producing recombinant Tir, and a kit for the detection of Tir unduce a protective immune response is also provided. In addition, Tir fusion proteins can be used in attenuated E. coli to induce a cell-mediated immune response to other polypeptides, continue in the the binding of bacterial pathogens to their polyperial pathogens to their
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New translocated intimin receptor useful for treating infection by enteropathogenic or enterohemorrhagic Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                   "given as Xaa in the specification; Lys is deduced from the DNA sequence"
                                                                                                                                                                   Tir; translocated intimin receptor; Hp90; enteropathogenic; EPEC; infection; dlagnosis; vaccine.
                                                                                                                                                                                                                                        Key Location/Qualifiers

Domain 734..253

Domain 764..386

Misc-difference 180
                                                                                                                                    EPEC E. coli translocated intimin receptor (Tir).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Stein M;
                                                                                                                                                                                                                                                                                                                                                                   /note= "encoded by AAA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                claim 6; Page 55-58; 91pp; English.
                               AAY06220 standard; Protein; 549 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kenny B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           receptors is further provided
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (UYBR-) UNIV BRITISH COLUMBIA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-CA01042.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97US-0065130.
                                                                                                      16-AUG-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Devinney R, Finlay BB,
                                                                                                                                                                                                                                                                                                                                                                                                          /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-337712/28.
                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               549 AA;
                                                                                                                                                                                                                             Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAX58858
                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09924576-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20-MAY-1999.
                                                                    AAY06220
RESULT
AAY06220
                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 20; Length 559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
```

Sequence

~

```
02-SEP-1998;
09-SEP-1998;
09-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                              08-AUG-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-SEP-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-SEP-1998;
01-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-SEP-1998;
02-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                             AAY99408;
                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                                                                                                                 312
                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                  AAY99408
                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                 ŏ
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Intimin; Tir binding domain; inhibitor; intimin adhesion; screening; Tir-independent eukaryotic cell binding activity; bacterial infection; diarrhoea; antibacterial.
                                                                                                                                                                                                            59 GDNRASDVPGLPVNPMRLAA--SEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDG 116
                                                                                  117 KHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
                                                                                                                 179 trilakldpdnhggrqpkdvdtrsvgvgsasgi----ddgv--vsethtsttnssvrs 230
                                                                                                                                                                                     408 VENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSS------IGGPCRIR 458
                                                                                                                                                                                                                                                                                               404 vvqqqtgg1pqhkvalmpqerrrfsdrrdsqgsvasthwsdsssevvnpyaevgg---ar 460
                                                                                                                                                                                                                                                                                                                459 MIMIKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPP----RDTTDNG---ARLLG 510
                                                                                                                                                                                                                                                                                                                            DPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEA 287
                                                                                                                                                                                                                                              QAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSART 407
                       70; Gaps
                                                 1 MPIGNLGHNPNVNNSIPPAAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS 58
                                                                                                                                           QRILELLEPKGTG-----ESKGAGESKGVGELRESNSGAENTTETQTSTSTSLRS
       Length 549;
                                                                                                                                                                                                                                                                                                                                                           511 NPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Intimin C-terminal Tir binding domain amino acid sequence.
                       Indels
      52.5%; Score 1498; DB 20;
54.8%; Pred. No. 6.6e-108;
ive 64; Mismatches 132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dougan G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hale CB,
                                                                                                                                                                                                                                                                                                                                                                                                             Ā
                                                                                                                                                                                                                                                                                                                                                                                                           AAB20576 standard; protein; 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-JAN-2000; 2000WO-GB00254
                                                                                                                                                                                                                                                                                                                                                                                                                                             08-DEC-2000 (first entry)
               Best Local Similarity 54.8%
Matches 323; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matthews SJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200045173-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Frankel GM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-AUG-2000
                                                                                                                                                                                                                                                                                                                                                                                                                             AAB20576;
       Query Match
                                                                                                                                                                                           231
                                                                                                                                                                                                             288
                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                               348
                                                                                                                                                                                                                                                                                                                                 461
                                                                                                                                                                                                                                                                                                                                                                                                    AAB20576
                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                        g
                                                                                                         õ
                                                                                                                        g
                                                                                                                                          ò
                                                                                                                                                          셤
                                                                                                                                                                           ò
                                                                                                                                                                                         g
                                                                                                                                                                                                           δ
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                δ
                                        õ
                                                      g
```

```
The present invention describes a method of screening for an inhibitor of intimin binding to eukaryotic cells. The method comprises exposing an intimin polypeptide having a Tir-independent cell binding activity to test agents, and obtaining an inhibitor based on its ability to bind the polypeptide. The inhibitors are used in the prevention, treatment and/or aliagnosis of bacterial infections, preferably by enteropathic and/or enterobaemorrhagic Escharichia coll, Shiga toxigenic E. coll, Hafnia alvei or Citrobacter freudil, Or especially E. coll ol5:Hy. The infections cause a histopathological effect known as attachment and effacement on intestinal epithelial cells. The inhibitors can be used to produce food supplements or additives, especially where the food is a milk substitute. The method can be used to sort cells based on their ability to bind to a Tir independent cell binding domain of an intimin polypeptide. Polypeptides having Tir-independent intimin binding cattle of the present sequence represents a specifically claimed intimin cerminal contributions. The present sequence represents a specifically claimed intimin cerminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLK 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human; PRO polypeptide; membrane bound protein; receptor; diagnosis; transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
                                                                   Screening for inhibitors of intimin binding to eukaryotic cells, use in diagnosing, preventing and treating bacterial infections, especially Escherichia coli 0157:H7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14.2%; Score 406; DB 21; Length 107; 72.0%; Pred. No. 2.9e-24; tive 18; Mismatches 12; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human PR01342 (UNQ697) amino acid sequence SEQ ID NO:243.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY99408 standard; Protein; 596 AA.
                                                                                                                                                                                                               Claim 8; Page 76; 96pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0098716.
98US-0098749.
98US-0098803.
98US-0098821.
98US-0098843.
98US-0098843.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-US20111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 72.04
Matches 77; Conservative
WPI; 2000-499357/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        107 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200012708-A2
```

```
98US-0099598.
98US-0099642.
98US-0099741.
98US-0099741.
98US-0099763.
98US-0099815.
98US-0099816.
98US-0099816.
98US-0100381.
98US-0100381.
98US-0100381.
98US-0100381.
98US-0100584.
98US-0100584.
98US-0100584.
98US-0100584.
98US-0100584.
98US-010041.
98US-010041.
98US-010041.
98US-010147.
                                                                                                                                                                                                                                                                                                                                                     980S-0102487
980S-0102570
980S-0102570
980S-0102684
980S-0102968
980S-0103258
980S-0103258
980S-0103314
980S-0103315
980S-0103315
980S-0103328
                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0103401.
98US-0103633.
98US-0103678.
98US-0103679.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0105002.
98US-0105104.
98US-0105169.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0104257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0104987
                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-071-1998

08-077-1998

08-077-1998

08-077-1998

14-077-1998

20-077-1998

20-077-1998

21-077-1998

21-077-1998
                                                                                                                                                                                                                                                                                                                                 29-SEP-1998
29-SEP-1998
30-SEP-1998
30-SEP-1998
30-SEP-1998
01-OCT-1998
02-OCT-1998
06-OCT-1998
                                                                                                                                                                                                                                                                                                            29-SEP-1998;
29-SEP-1998;
29-SEP-1998;
                                                                                                                                                                                                                                                                                        24-SEP-1998;
24-SEP-1998;
24-SEP-1998;
                                                                              15-SEP-1998;
15-SEP-1998;
15-SEP-1998;
16-SEP-1998;
16-SEP-1998;
16-SEP-1998;
                                                                                                                                                                                                            -SEP-1998;
-SEP-1998;
-SEP-1998;
                                                                                                                       16-SEP-1998;
16-SEP-1998;
17-SEP-1998;
                                                                                                                                                          17-SEP-1998;
17-SEP-1998;
17-SEP-1998;
18-SEP-1998;
                                                                                                                                                                                                                                                              23-SEP-1998
                                                                                                                                                                                                                                                                                   -SEP-1998
```

```
receptor or secreted PRO polypeptides given in AAY99340 to AAA94462. The receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoling then have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridisation probes used in the isolation of the PRO polypeptides from the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides, useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Wood WI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Watanabe CK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Smith V,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 12; Fig 138; 773pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gurney AL,
                              9805-0105693.
98085-0105694.
98085-0105807.
98085-0105882.
98085-01060623.
98085-0106023.
98085-0106033.
98085-0106033.
98085-0106033.
98085-0106178.
98085-0106178.
98085-0106178.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9805-0108848-9805-0108849-9805-0108849-9805-0108851-9805-0108851-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108898-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-0108858-9805-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 980S-0107783.
980S-0108775.
980S-0108779.
980S-0108787.
980S-0108788.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0108802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0108867
      98US-0105266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Baker K, Goddard A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-237871/20.
N-PSDB; AAA37090.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            596 AA;
22-OCT-1998; 26-OCT-1998; 27-OCT-1998; 27-OCT-1998; 27-OCT-1998; 27-OCT-1998; 28-OCT-1998; 28-OCT-1998; 28-OCT-1998; 28-OCT-1998; 28-OCT-1998; 29-OCT-1998; 29-OCT-1998; 30-OCT-1998; 30-OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
```

ò

Query Match 5.2%; Score 148.5; DB 21; Length 596; Best Local Similarity 17.7%; Pred. No. 0.0029; Matches 102; Conservative 84; Mismatches 255; Indels 135;

19;

Gaps

```
QRNGVETS--VVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                      : : | | | : : :||| | : : :||| | : :||| :|| | : :||| | : | | | | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : :||| | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                    241 GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       n---sesstvssrastatnsesstt---ssgastatnsesrttsngagtatnsesstts 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QEELKVSSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVEN 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sessttssgastatnsdssttssgagtatnsesstvssgistvtnsesstpssgantatn 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRR 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CRLLILIRIFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLA----- 524
                                                                                    SDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV-FRVETQEDGKHIAVG 122
                                                                                                                                                                                                                                                            sessttssgastatnse---sstpssgastvtnsgssvtssgastatnsesstvssras 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301 LG------NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sgastatnsdsstvssga---statnsessttssgast-------atn 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----vstatnsessttssgastatnsdssttsseastatnsess 441
                                                                                                                                          ssvtssgvstatisgssvtsn-gvsiv-tnsefhttssgistatnsefstassgislatn 107
      | : |
-----astatnsg
                                                                                                                                                                                                                                                                                                                  181 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -LSGGL----RHDMGGLTGGSNSAVNTSNNPPAPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB87575 standard; Protein; 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20000S-0175481.
2000WO-US04341.
2000WO-US04342.
2000WO-US0414.
2000WO-US05601.
2000US-0187202.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-US21090.
99US-0169495.
99US-0170262.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000WO-US14042
2000US-0209832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000WO-US23328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99WO-US20111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human; PRO protein; mapping
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200116318-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human PRO1342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-SEP-1999;
07-DEC-1999;
09-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-FEB-2000;
22-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-2000;
03-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22-MAY-2000;
05-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-APR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB87575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         442
                                                                                                                                                                                                                                                         801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB87575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                          硆
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ·δ
```

```
The present sequence is a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ssvtssgvstatisgssvtsn-gvsiv-tnsefhttssgistatnsefstassgisiatn 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QRNGVETS--VVLSDQEYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSTSSLRSDPKLWLALGTVAT 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sgastatnsdsstvssga---statnsessttssgast-----atn 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     352 QEELKVSSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVEN 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sessttssgastatnsdssttssgagtatnsesstvssgistvtnsesstpssgantatn 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRR 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV-FRVETQEDGKHIAVG 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sessttssgastatnse---sstpssgastvtnsgssvtssgastatnsesstvssras 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tatnsesstlssgastatn-sdssttssgastatnsessttssg-----astat 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            n----sesstyssrastatnsesstt---ssgastatnsesrttsngagtatnsesstts 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------vstatnsessttssgastatnsdssttsseastatnsess 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301 LG------NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNRA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13 glilhleaatns-----astatnsg 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLA-----
                                                                                                                                                                            Eighty four nucleic acids encoding PRO polypeptides, useful in molecular biology, including use as hybridization probes, and in
                                                        Godowski PJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.2%; Score 148.5; DB 22; Length 596; 17.7%; Pred. No. 0.0029; Live 84; Mismatches 255; Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         363 sessttssgantatnsesstvssgastatnsess---ttssg-------
                                                        Goddard A,
Wood WI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -LSGGL----RHDMGGLTGGSNSAVNTSNNPPAPGS 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tvssgistvtnsessttssgantatnsgssvtsags 477
                                                        S, X
                                                        Gerritsen Natanabe (
                                                                                                                                                                                                                                                         Claim 12; Fig 100; 278pp; English.
                                                                                                                                                                                                                    chromosome and gene mapping
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 5.2%
Best Local Similarity 17.7%
Matches 102; Conservative
                                                        Ē,
ĀĽ,
                (GETH ) GENENTECH INC
                                                        Filvaroff
                                                                                                                  WPI; 2001-183260/18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               596 AA;
                                                                                                                                     N-PSDB; AAF92107
                                                                            Grimaldi CJ,
                                                        DL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                          Eaton
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       442
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

RESULT

us-09-189-415a-4.rag

AAB66157;

AAB66157

```
411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRR 470
                                                                                                                                                                                                                                                                                                                                                                            ----vstatnsessttssgastatnsdssttsseastatnsess 441
                                                                                                                                                                                                352 QEELKVSSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTTVEN 410
tatnsesstlssgastatn-sdssttssgastatnsessttssg-----astat 211
                                                    GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                       n----sesstvssrastatnsesstt---ssgastatnsesrttsngagtatnsesstts
                                                                                                                           301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKR
                                                                                                                                                 sessttssgastatnsdssttssgagtatnsesstvssgistvtnsesstpssgantatn
                                                                                                                                                                                                                                                                                                                                             471 CRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The mutant protease having mew cleavage specificities is obtained by deleting three amino acids. This mutant may then be used to prepare hybrid proteases,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Venema G,
                                                                                                                                                                                                                                                                                                           363 sessttssgantatnsesstvssgastatnsess---ttssg----
                                                                                                                                                                                                                                                                                                                                                                                                                  525 -LSGGL----RHDMGGLTGGSNSAVNTSNNPPAPGS 555
                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 1(1-7)+5(b); 29 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kok J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label= signal_peptide
188..1959
/label= mature_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     De Vos WM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR10562 standard; Protein; 1959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mutant protease (delta137-139)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89NL-0002010.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90EP-0202113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lactococcus lactis SK11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Vos PAJ, Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1991-038622/06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAQ10416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                        402 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EP411715-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR10562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide
                          164
                                                                                                                                                                                                                                             303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR10562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                    δ
                                                                                                                                                                      q
                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Secreted and transmembrane proteins and nucleic acids designated PRO, useful as hybridization probes, in chromosome and gene mapping and gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fong S;
Hillan KJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present invention relates to secreted and transmembrane proteins. These proteins and the DNA encoding them may be used as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. They may also be used used to generate either transgenic animals or knockout animals which are in turn useful for development and screening of therapeutically useful reagents. The nucleic acids may also be used in gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVAT 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV-FRVETQEDGKHIAVG 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ssvtssgvstatisgssvtsn-gvsiv-tnsefhttssgistatnsefstassgisiatn 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QRNGVETS--VVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     108 sessttssgastatnse---sstpssgastvtnsgssvtssgastatnsesstvssras 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 GNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNRA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           glllhleaatns----astatnsg 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.2%; Score 148.5; DB 22; Length 596;
7 17.7%; Pred. No. 0.0029;
vative 84; Mismatches 255; Indels 135;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ferrara N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D; Watanabe CK, Williams PM, Wood WI;
                                                                                                                                                                               Secreted; transmembrane; gene therapy
                               AAB66157 standard; protein; 596 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       claim 1; Fig 138; 787pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                               990S-0162506.
99WO-US28313.
99WO-US28551.
99WO-US30095.
2000WO-US00219.
2000WO-US00376.
                                                                                                                                                                                                                                                                                                                                                          99US-0141037.
99US-0144758.
99US-0145698.
                                                                                                                                            Protein of the invention #69
                                                                                                                                                                                                                                                                                                                            2000WO-US04342
                                                                                                                                                                                                                                                                                                                                                                                                                  99WO-US20111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-071395/08
                                                                                                          (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       596 AA;
                                                                                                                                                                                                                                                      WO200078961-A1
                                                                                                                                                                                                                     Unidentified.
                                                                                                                                                                                                                                                                                                                            18-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16-DEC-1999;
05-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Sim
hes 102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-NOV-1999
02-DEC-1999
                                                                                                                                                                                                                                                                                                                                                               23-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                20-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                    01-SEP-1999
                                                                                                            02-APR-2001
                                                                                                                                                                                                                                                                                          28-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                   29-0CT-1999
```

Haandrikman AJ;

Seguence

therapy

SACANA SA

Query Match

Best Loca Matches

13

ò q à 셤 ò

20 64

123

g

181

å

```
Similarity
                                                                                                                                                                                            N-PSDB; AAQ10414.
                                                                                02-AUG-1990;
                                                                                                          04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Simi
Matches 114;
                                                          06-FEB-1991
                                   EP411715-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1510
                        合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                          20;
                                                                                                                                                                                                                                                                                                                                slktkvaaaveaaktvgkgdgttgtsdkggggggtpapapgdtgkdkgdegsgpssggnip 1836
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : |||:||| |: | | : | |:
tkpatttstttddttdrngqltsgtsdkgggggtpapapgdigkdkgdegsgpsgggnip 1896
                                                                                                                                                                                                                                -----ttepaqtvtltanaaa----tgetvqysadggktyqdvp-----aagvtit 1548
                                                                                                                                                                                                                                                                                                                                                                           | : | : | | | | | atttalaaatqkaq-taldqtnasvdsltganrdlqtainqlaaklpadkktslinqlqs 1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1717 ldevlaklaegikaatpaevgnakdaatgktwyadiadtltsgqasadasdklahlgalq 1776
                                                                                                                                                                                                                                                                 GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                198 KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                               TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                            311 KDDVVANIEEQAKAAGE----EAKQQAI------ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          390 VEQTITITITITISAR-----TVENKPANNIPAQGNVD---TPGSEDIMESRRSSM- 438
                                                                                                                                                                 ; Score 141.5; DB 12; Length 1959;
; Pred. No. 0.051;
61; Mismatches 211; Indels 113; Gaps
                                                                                                                                                                                                                 36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                          The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis et compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

See also AAQ10411-17 and AAQ10870-71.
the fusion being between a type I and a type III protease of L.lactis Wg2 and SK11. The product has modified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mutant protease gene; fermentation; foodstuff; flavouring;
lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352 QEELKVSS-----GAGYGLSGALILGGGIGVAVTA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         188..1962
/label- mature_protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1.187
/label- signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR10560 standard; Protein; 1962 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                 5.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :||:| ||: |
1897 tnpatttstttddttdrng 1915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                              Best_Local Similarity 22.8
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lactococcus lactis SK11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mutant protease (K138N)
                                                                                                                              1959 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-APR-1991
                                                                                                                                Sequence
                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR10560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Protein
                                                                                                                                                                                                                                        1449
                                                                                                                                                                                                                                                                86
                                                                                                                                                                                                                                                                                                                                        1549
                                                                                                                                                                                                                                                                                                                                                                                       1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          439
                                                                                                                                                                                                                                                                                        1507
                                                                                                                                                                                                                                                                                                                                                                                                               257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
AAR10560
ID AAR10560
XX
AC AAR1
XX
DT 15-AI
XX
KW Mutai
XX
KW Mutai
XX
COS Lacti
XX
COS Lacti
FT Pept
FT Pept
                      8888888888
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1612 atttalaaatqkaq-taldqtnasvdsltganrdlqtainglaaklpadkktsllnglgs 1670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1720 ldevlaklaegikaatpaevgnakdaatgktwyadiadtltsgqasadasdklahlgalq 1779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----ttepaqtvtltanaaa----tgetvqysadggktyqdvp------aagvtit 1551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QEELKVSS------GAGYGLSGALILGGGIGVAVTA------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The mutant protease K138N having mew cleavage specificities is obtained by carrying out single amino acid substitutions. This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of L.lactis Wg2 and SK11.

The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis et compared to the parent protease(s). The proteases can be used for preparing products (butter ofbese, human and animal foodstuffs) see also AAQ10411-17 and AAQ10870-71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             390 VEQTTTTTTTTTTTSR-----TVENKPANNTPAQGNVD---TPGSEDTMESRRSSM-
                                                                                                                                                                                                                                                                                                                               from type I and
                                                                                                                                                                                                                                                                                                                               Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 12; Length 1962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      211;
                                                                                                                                                                                Venema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.0%; Score 141.5; DF
22.8%; Pred. No. 0.051;
iive 61; Mismatches ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Fig 1(1-7)+5(a); 29 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    311 KDDVVANIEEQAKAAGE----EAKQQAI-----
                                                                                                                                                                             Kok J,
                                                                                                                                                                                ¥,
                                                                                                                                                                                De Vos
90EP-0202113
                                                          89NL-0002010
                                                                                                                     NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                              III protease genes from lac
foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                Vos PAJ, Siezen RJ,
                                                                                                                                                                                                                                     WPI; 1991-038622/06.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1962 AA;
```

<u>م</u> م

```
1840 tkpatttstttddttdrnggltsgtsdkggggggtpapapgdigkdkgdegsgpssggnip 1899
                                                                                               .612 atttalaaatqkaq-taldqtnasvdsltganrdlqtainqlaaklpadkktsllnqlqs 1670
                                                                                                                                                                                                                       : ||::
1780 slktkvaaaveaaktvgkgdgttgtsdkggggggtpapapgdtgkdkgdegsgpssggnip 1839
                                                                                                                                                         1552 angtíkístálygnespavdyvvtnikaddpaqlqaakqeltnliasaktisasgkydd 1611
                                                                                                                                                                                                                                                                   352 QEELKVSS------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                     390 VEQITITITITITISAR-----TVENKPANNIPAQGNVD---IPGSEDIMESRRSSM- 438
                                                                                                                                  257 TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The mutant protease Al37G/Kl38D having mew cleavage specificities is
                                                                 198 KGVGELRESNSGAENTTETQTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                    ----ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
 ---TEARQRILELLEPKGTGESKGAGES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Venema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Fig 1(1-7)+5(b); 29 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kok J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1..187
/label= signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     188..1962
/label= mature_protein
                                                                                                                                                                                                         311 KDDVVANIEEQAKAAGE----EAKOQAI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR10557 standard; Protein; 1962 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
       146 PEGKDKFVFTGGRGGAGHAMVTVASDI --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         De Vos WM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mutant protease (A137G/K138D)
                                                                                                                                                                                                                                                                                                                                                                                                                                :||:| ||: |
1900 tnpatttstttddttdrng 1918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90EP-0202113.
                                                                                                                                                                                                                                                                                                                                                                                                               439 ---- ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          89NL-0002010.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lactococcus lactis SK11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jos PAJ, Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1991-038622/06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAQ10411.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP411715-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR10557;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                        δ
                                                                                                        g
                                                                                                                                         Q
                                                                                                                                                                       g
                                                                                                                                                                                                            à
                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1452 insgkp-ghmaidgpvkllegknvltvavtdsedntttknitvyyepkktlaaptvtps- 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis etc., compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria. See also AAQ10411-17 and AAQ10870-71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                86 GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The mutant protease N166D having mew cleavage specificities is obtained by carrying out single amino acid substitutions. This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of L.lactis Mg2 and SK11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 12; Length 1962;
                                                                                                                                                                                                                                                Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Venema G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 141.5; DB red. No. 0.051; 61; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Fig 1(1-7)+5(a); 29 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kok J,
                                                                                                                                                                                                                                                                                                                                                      1..187
/label= signal_peptide
188..1962
/label= mature_protein
                                                                                                                  AAR10561 standard; Protein; 1962 AA.
                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              De Vos
                               1900 tnpatttstttddttdrng 1918
----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90EP-0202113.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89NL-0002010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (NEZU-) NED INST ZUIVELONDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 5.0°
Best Local Similarity 22.8°
Matches 114; Conservative
                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                     Lactococcus lactis SK11
                                                                                                                                                                                                                  Mutant protease (N166D)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1991-038622/06.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1962 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAQ10415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-FEB-1991
                                                                                                                                                                                  15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                         EP411715-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Vos PAJ,
                                                                                                                                                    AAR10561;
                                                                                                                                                                                                                                                                                                                                          Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                           Protein
 439
```

ద

g S

σ

```
Vos PAJ, Siezen RJ,
                                                                                                                                                                            WPI; 1991-038622/06
                                                                                                                                                                                                                                                                                                                                                                                                                           1962 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                        N-PSDB; AAQ10412.
                                                                                 02-AUG-1990;
                                                                                                        04-AUG-1989;
                                                           06-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114;
                                   EP411715-A.
                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
   Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                        20;
                                                                                                                                                                                                                                                                                                                                                                                                                 1452 insgkp-ghmaidgpvkllegknvltvavtdsedntttknitvyyepkktlaaptvtps- 1509
                                                                                                                                                                                                                                                                         -----ttepaqtvtltanaaa----tgetvqysadggktyqdvp-----aagvtit 1551
                                                                                                                                                                                                                                                                                                                       atttalaaatqkaq-taldqtnasvdsltganrdlqtainqlaaklpadkktsllnqlqs 1670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1780 slktkvaaaveaaktvgkgdgttgtsdkgggggtpapapgdtgkdkgdegsqpssggnip 1839
                                                                                                                                                                                                                                                                                                                                                                                                   257 TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QEELKVSS-----ALHRKNQP 389
                                                                                                                                                                                                                                                           86 GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                        146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                     198 KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKIWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                  311 KDDVVANIEEQAKAAGE----EAKQQAI-------ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            390 VEQTITITITITITSAR----TVENKPANNIPAQGNVD---TPGSEDIMESRRSSM- 438
                                                                                                                                                                                       Indels 113; Gaps
                                                                                                                                                                                                             36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
obtained by replacing two amino acids.

This mutant may then be used to prepare hybrid proteases, this mutant may then be used to prepare hybrid protease of L.lactis Mg2 and SK11.

L.lactis Mg2 and SK11.

The product has modified properties, e.g. thermostability, alloprocessed to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

See also AQ10411-17 and AAQ10870-71.
                                                                                                                                                                DB 12; Length 1962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protease gene; fermentation; foodstuff; flavouring;
                                                                                                                                                               5.0%; Score 141.5; DB 12;
22.8%; Pred. No. 0.051;
tive 61; Mismatches 211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1..187
/label- signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mutant protease (A137G/K138P/T139P)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR10558 standard; Protein; 1962 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :||:| ||: |
1900 tnpatttstttddttdrng 1918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                           Best_Local Similarity 22.8
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lactococcus lactis SK11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mutant protease gene;
lactic acid bacteria.
                                                                                                                              1962 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-APR-1991
                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR10558
                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
                                                                                                                                                                                                                                                                                 1510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88888888888888888
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
| || | | | | : | : |
1552 angtfkfkstdlygnespavdyvvtnikaddpaqlqaakqeltnliasaktlsasgkydd 1611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1510 -----ttepaqtvtltanaaa----tgetvqysadggktyqdvp-----aagvtit 1551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1780 slktkvaaaveaaktvgkgdgttgtsdkgggggtpapapgdtgkdkgdegsgpssggnip 1839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           146 PECKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              198 KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257 TPE---POSPTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          311 KDDVVANIEEQAKAAGE----EAKQQAI------EONAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   390 VEQTITITITITISAR----TVENKPANNIPAQGNVD---TPGSEDIMESRRSSM- 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                                                                                                                                                                                                                                                                              Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The mutant protease having mew cleavage specificities is obtained by replacing three amino acids.
This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of L.lactis Wg2 and SKII.
The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.
See also AAO10411-17 and AAO10870-71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.0%; Score 141.5; DB 12; Length 1962; 22.8%; Pred. No. 0.051; vative 61; Mismatches 211; Indels 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                              o,
                                                                                                                                                                                                                                                                                                                                                                                                                                              Venema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QEELKVSS-----GAGYGLSGALILGGGIGVAVTA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Disclosure; Fig 1(1-7)+5(b); 29 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                              Kok J,
188..1962
/label= mature_protein
                                                                                                                                                                                                                                                                                                                                                                                                                                              MΨ
                                                                                                                                                                                                                                                                                                                                                                                                                                              De Vos
                                                                                                                                                                                                                                                                                                        89NL-0002010
                                                                                                                                                                                                                                     90EP-0202113
                                                                                                                                                                                                                                                                                                                                                                       (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
```

Thu

```
Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mutant |
lactic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein
        1510
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                           q
                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                              à
                                                                           δ
                                                                                                셤
                                                                                                                           δ
                                                                                                                                             Q
                                                                                                                                                                     Ω
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 obtained by replacing three amino acids.
This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of L.lactis Wg2 and SKII.

The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis etc., compared to the parent proteases(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

See also AAQ10411-17 and AAQ10870-71.
1840 tkpatttstttddttdrnggltsgtsdkgggggtpapapgdigkdkgdegsqpssggnip 1899
                                                                                                                                                                                                                                                                                                                                                                                                                                            Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    from type I and in fermentation
                                                                                                                                                                                               Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        mutant protease having mew cleavage specificities is
                                                                                                                                                                                                                                                                                                                                                                                                                                            Venema G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mutant protease gene(s) and protease(s) - derived
III protease genes from lactococcal strains, used
foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 1(1-7)+5(b); 29 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                             ,
,
                                                                                                                                                                                                                                                                   1.187
/label= signal_peptide
188.1962
/label= mature_protein
                                                                                                                                                                                                                                                                                                                                                                                                                                             Kok
                                                                                                       AAR10559 standard; Protein; 1962 AA.
                                                                                                                                                                           Mutant protease (A137G/K138L/T139A).
                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                              WW.
                                     :||:| ||: |
1900 tnpatttstttddttdrng 1918
                                                                                                                                                                                                                                                                                                                                                                                                                                               De Vos
                        439 ----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                          90EP-0202113
                                                                                                                                                                                                                                                                                                                                                                                                89NL-0002010
                                                                                                                                                                                                                                                                                                                                                                                                                       (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                     Lactococcus lactis SK11.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1991-038622/06.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1962 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAQ10413
                                                                                                                                                                                                                                                                                                                                                                          02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                 04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                     06-FEB-1991
                                                                                                                                                      15-APR-1991
                                                                                                                                                                                                                                                                                                                               EP411715-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                               Vos PAJ,
                                                                                                                                AAR10559;
                                                                                                                                                                                                                                                                     Pept1de
                                                                                                                                                                                                                                                                                              Protein
                                                                                                                     ò
```

```
: |||:||| |: | |: | ::
1840 tkpatttstttddttdrngqltsgtsdkgggggtpapapgdigkdkgdegsgpssggnip 1899
                                                                                                                                                                                                                                                                             1612 atttalaaatgkag-taldgtnasydsltganrdlgtainglaaklpadkktslinglgs 1670
                                                                                                                                                                                                                                                                                                                                                                                                                       1671 vkdalgtdlgngtdp--stgktftaalddlvaqaqagtqtddqlgatlak------1 1719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1720 ldevlaklaegikaatpaevgnakdaatgktwyadiadtltsgqasadasdklahlqalq 1779
                                                                                                                                          | | : | | : | | : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                            257 TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      311 KDDVVANIEEQAKAAGE----EAKQQAI-------EUNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         352 QEELKVSS-----AGAGYGLSGALILGGGIGVAVTA------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 VEQTTTTTTTTTTTTSAR-----TVENKPANNTPAQGNVD---TPGSEDTMESRRSSM- 438
                                                                                                                                                                                                                                     198 KGVGELRESNSGAENTTETQTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                              146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1780 slktkvaaaveaaktvgkgdgttgtsdkggggggtpapapgdtgkdkgdegsgpssggnlp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene; fermentation; foodstuff; flavouring;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              o,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Venema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Fig 1(1-7)+5(a); 29 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kok J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /label= signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /label= mature_protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR10563 standard; Protein; 1962 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥Μ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :||:| ||: |
1900 tnpatttstttddttdrng 1918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      De Vos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          90EP-0202113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89NL-0002010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..1962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         439 ----ASTSSTFFDTSSIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mutant protease (K748T)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lactococcus lactis SK11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protease gene;
acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Vos PAJ, Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1991-038622/06.
N-PSDB; AAQ10417.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   04-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EP411715-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR10563;
```

20;

5.0%; Score 141.5; DB 12; Length 1962; 22.8%; Pred. No. 0.051; tive 61; Mismatches 211; Indels 113; Gaps

Best Local Similarity 22.8 Matches 114; Conservative

Query Match

셤

à

36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85

us-09-189-415a-4.rag

```
-----ttepaqtvtltanaaa----tgetvqysadggktyqdvp-----aagvtit 1557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | || | | | | : | : |
1558 angtfkfkstdlygnespavdyvvtnikaddpaqlqaakqeltnliasaktlsasgkydd 1617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1618 atttalaaatgkag-taldgtnasvdsltganrdlgtainglaaklpadkktsllnglgs 1676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1677 vkdalgtdlgngtdp--stgktftaalddlvagagagtgtddglgatlak-----1 1725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1726 ldevlaklaegikaatpaevgnakdaatgktwyadiadtltsgqasadasdklahlgalg 1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        352 QEELKVSS------GAGYGLSGALILGGGIGVAVTA------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                                                                                 Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteolysis et compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria. See also AAQ10411-17 and AAQ10870-71.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --- ENNAQAQKKYDEQQAKR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       198 KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL
                                                                                                                                                                                                                                                                                                                                  Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                                                                                                                                                                                                                                                                                                                                                               The wild-type L.lactis SK11 protease gene sequence was determined by the applicant (EP-307011).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The mutant protease having mew cleavage specificities is obtained by deleting three amino acids (nine bps) and inserting
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease L.lactis Wg2 and SX11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                 Venema G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        311 KDDVVANIEEQAKAAGE----EAKQQAI-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.0%; Score 141.5; DE 22.8%; Pred. No. 0.052; iive 61; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 1(1-7)+5(b); 29pp; English
                                                                                                                                                                                                                                                   ٦,
                                                                                                                                                                                                                                                 Kok
             /label= sig_peptide
188..1968
/label= mat_protein
                                                                                                                                                                                                                                                   WW,
                                                                                                                                                                                                                                                   De Vos
                                                                                                                                                  90EP-0202113.
                                                                                                                                                                                 89NL-0002010
                                                                                                                                                                                                                  (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                 Vos PAJ, Siezen RJ,
                                                                                                                                                                                                                                                                                  WPI; 1991-038622/06.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
les 114; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9 other residues.
                                                                                                                                                                                                                                                                                                   N-PSDB; AAQ10871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1968
                                                                                                                                                02-AUG-1990;
                                                                                                                                                                               04-AUG-1989;
                                                                                                                 06-FEB-1991
                                                                                 EP411715-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
 Peptide
                                  Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1516
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1780 slktkvaaaveaaktvgkgdgttgtsdkgggggtpapapgdtgkdkgdegsgpssggnip 1839
                                                                                                                                                                                                                                                                                                                                                       1671 vkdalgtdlgngtdp--stgktftaalddlvaqaqagtqtddqlqatlak-----1 1719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1510 -----ttepaqtvtltanaaa----tgetvqysadggktyqdvp-----aagvtit 1551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        198 KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 257 TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   352 QEELKVSS-----GAGYGLSGALILGGGIGVAVTA------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 VEQTITITITITITITAR-----TVENKPANNTPAQGNVD---TPGSEDIMESRRSSM- 438
                                                                                                                                                                                                                                                                                                                                                                                                  86 GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 311 KDDVVANIEEQAKAAGE----EAKQQAI------ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146 PEGKDKFVFTGGRGGAGHAMYTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                  36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                             obtained by carrying out single amino acid substitutions.
This mutant may then be used to prepare hybrid proteases,
the fusion being between a type I and a type III protease of
L.lactis Wg2 and SK11.
The product has modified properties, e.g. thermostability,
alkaline/acid pH stability, oxidative stability, autoproteolysis et
compared to the parent protease(s). The proteases can be used for
prepared with the aid of lactic acid bacteria.
See also AAQ10411-17 and AAQ10870-71.
                                                                                                                                                                                                                                                                  DB 12; Length 1962;
                 ţ
                 mutant protease K748T having mew cleavage specificities
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protease gene; fermentation; foodstuff; flavouring; acid bacteria.
                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                               211;
                                                                                                                                                                                                                                                               5.0%; Score 141.5; Di 22.8%; Pred. No. 0.051
                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mutant protease (delta137-139/ins9)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR10941 standard; Protein; 1968
                                                                                                                                                                                                                                                                                                 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :||:| ||: |
1900 tnpatttstttddttdrng 1918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      439 ----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-APR-1991 (first entry)
                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lactococcus lactis SK11
                                                                                                                                                                                                                1962 AA;
                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                   Best Local Sim
Matches 114;
                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR10941;
                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mutant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    *****************
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

20;

Gaps

```
셤
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                   οy
                                                        g
                                                                                                      qq
                                                                                                                                  à
                                                                                                                                                       g
                                                                                                                                                                                οy
                                                                                                                                                                                                                                  ò
                                   δ
                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15 other residues.
This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of L.lactis Mg2 and SK11.
The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autoproteelysis etc., compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.
See also AAQ10411-17 and AAQ10870-71.
                                                 1846 tkpatttstttddttdrnggltsgtsdkgggggtpapapgdigkddegsggsgggnip 1905
 | | | : ||::
1786 slktkvaaaveaaktvgkgdgttgtsdkgggggtpapapgdtgkdkgdegsgpssggnip 1845
                                   390 VEQTTTTTTTTTTSAR-----TVENKPANNTPAQGNVD----TPGSEDIMESRRSSM- 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Haandrikman AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The wild-type L.lactis SKII protease gene sequence was determined by the applicant (EP-307011).

The mutant protease having maw cleavage specificities is obtained by deleting three amino acids (nine bps) and inserting
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings
                                                                                                                                                                                                                                                                      Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ο,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Venema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig 1(1-7)+5(b); 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kok J,
_
=
                                                                                                                                                                                                                                                                                                                                     1..187
/label= SIG_PEPTIDE
188..1974
                                                                                                                                                                                                                                               Mutant protease (deltal37-139/ins15)
                                                                                                                                                                          AAR10940 standard; Protein; 1974 AA.
                                                                                                                                                                                                                                                                                                                                                                          188..1974
/label= MAT_PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        De Vos WM,
                                                                                                      :||:| ||: |
1906 tnpatttstttddttdrng 1924
                                                                                     439 ----ASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90EP-0202113.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        89NL-0002010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (NEZU-) NED INST ZUIVELONDE
                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                              lactis SK11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vos PAJ, Siezen RJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1991-038622/06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAQ10870.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        04-AUG-1989;
                                                                                                                                                                                                                                                                                                                Lactococcus
                                                                                                                                                                                                                         15-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                          06-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                 EP411715-A.
                                                                                                                                                                                                  AAR10940
                                                                                                                                                                                                                                                                                                                                                                            Protein
                                                                                                                                                                                                                                                                                                                                                    Peptide
                                                                                                                                                                            g
                                                                                                         q
                                                                                      à
```

DB 12; Length 1974;

Score 141.5;

5.0%;

Query Match

1974 AA;

Sequence

```
20;
                                                                                                                                                                                                            1683 vkdalgtdlgngtdp--stgktftaalddlvaqaqagtqtddqlqatlak-----1 1731
                                                                                                                                                                                                                                                                                                                                                                                                     1464 insgkp-ghmaidgpvkllegknvltvavtdsedntttknitvyyepkktlaaptvtps- 1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                   389
                                                                                                                                                                                                                                                                                                                    TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                        GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                     198 KGVGELRESNSGAENTTETQTSTSTSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                  311 KDDVVANIEEQAKAAGE----EAKQQAI------ENNAQAQKKYDEQQAKR 351
                            Gaps
                                                           ----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                      146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES
                                                                                                                                                                                                                                                                                                                                                                                                                                                   --ALHRKNOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 390 VEQTTTTTTTTTTSAR-----TVENKPANNTPAQGNVD---TPGSEDTMESRRSSM-
                            Indels
Pred. No. 0.052;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    352 QEELKVSS-----GAGYGLSGALILGGGIGVAVTA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search completed: September 27, 2001, 14:22:28
Job time: 124 sec
 22.8%; Fig.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :||:| ||: |
tnpatttstttddttdrng 1930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   439 ----ASTSSTFFDTSSIGG 453
                                                                INSTGPLGSRALFTPVR---
               Best Local Similarity 22.8
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1912
                                                                                                                                                               1522
                                                                                                                                                                                                                                                                                                                            257
                                                                36
                                                                                                                              98
```

OTHER PAGE BLANN NUSPTON

```
CITY: Hawthorne
STATE: New York
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-186-222-2
119
118
118
117.5
117.5
117.5
117.5
115.5
115.5
115.5
115.5
115.5
115.5
115.5
115.5
115.5
115.5
RESULT
                                                                                                                                                                                                                               (without alignments)
542.925 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 5, Al
Sequence 10,
Sequence 10,
Sequence 115,
Sequence 2, Al
Sequence 2, Al
Sequence 2, Al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, A Sequence 5, A Sequence 10, Sequence 10, Sequence 10, Sequence 5, A Sequence 5, A Sequence 5, A Sequence 5, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2, A Sequence 2, A Sequence 32, Sequence 32,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 6, 1
Sequence 6, 1
Sequence 5, 1
                                                                                                                                                                                                       ; Search time 21.2 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                 US-09-189-415A-4
2851
1 MPIGNLGHNPNVNNSIPPAP.....SNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence Sequence S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        197339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /cgn2_6/ptodata/2/laa/5A_COMB.pep:*
/cgn2_6/ptodata/2/laa/5B_COMB.pep:*
/cgn2_6/ptodata/2/laa/6A_COMB.pep:*
/cgn2_6/ptodata/2/laa/6B_COMB.pep:*
/cgn2_6/ptodata/2/laa/pcTUS_COMB.pep:*
/cgn2_6/ptodata/2/laa/pcTUS_COMB.pep:*
                           GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-617-697-10
US-08-409-995-5
US-08-88-467-5
US-08-913-942-5
US-08-913-942-2
US-08-913-942-2
US-08-913-942-2
US-08-913-942-2
US-08-913-942-2
US-08-913-942-2
US-08-913-942-2
US-08-913-172-15
US-08-393-703-5
US-08-393-703-5
US-08-386-495-10
PCT-US96-02331-15
PCT-US96-02331-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-10661A-5
US-08-728-470-10
US-08-719-641-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-038-682-2
US-08-302-832-2
US-08-530-198-2
US-08-728-470-2
US-08-617-697-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197339 segs, 20590346 residues
                                                                                                                                                                                                 September 27, 2001, 14:21:04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                              - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           115336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
15336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      124.5
124.5
124.5
123
                                                                                                                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                 OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                                                                                                                       :
o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              No.
```

```
16;
                                                                                                                                                                                                                                         20, Appl
6, Appli
                                                                   Appli
Appli
                                                                                                                                                                                                                                                                                                         Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 33,
                               sednence sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 461;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: USA

ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/186,222
FILING DATE: US/08/186,222
FILING DATE: US/08/186,222
RICHARD APPLICATION DATA:
APPLICATION NUMBER: US/08/186,225
FILING DATE: 19-MAR-1991
APPLICATION NUMBER: GB 9006400.7
FILING DATE: 22-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Villamizar, JOANN
REGISTRATION NUMBER: 4-17994/A
TELECOMMUNICATION INMBER: 4-17994/A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4.4%; Score 125.5; DB 1;
19.0%; Pred. No. 0.0051;
tive 70; Mismatches 160;
                                                        US-08-769-309A-5
US-08-994-570-5
US-08-59-397A-31
US-08-614-377A-7
US-08-142-648B-7
US-08-193-855-20
PCT-US96-03916-66
US-08-719-641-9
US-08-179-641-9
US-08-127-499A-26
US-08-127-498A-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-377-155-33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08186222
Patent No. 555907
GENERAL INFORMATION:
APPLICANT: Suri, Bruno
APPLICANT: Schmitz, Albert
TITLE OF INVENTION: Bacterial Vectors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ANDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: (914)785-7121
TELEFAX: (914)347-5769
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 461 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 19.08
Matches 84; Conservative
11536
11581
1780
1780
1026
1026
1026
1338
1338
1599
484
484
484
631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-08-186-222-2
```

```
RESULT 3
PCT-US95-10661A-5
                                                                           87;
                                   Query Match
Best Local
                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                      염
                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     336 IGGSGINSSPIGNPYAVGGCTDYVWQYFAAQGIYIRNIMPGNGGQWASNGPAQGVLHVVG 395
                                                                                                                                                                                          325 AGEEAKQQAIENNAQAQKKYDEQQAKRQE--ELKVSSGAGYGLSGALILGGGIGVAVTAA 382
                                                                                                                                                                                                                                                                                                                      383 LHRKNQPVEQTITITITITISARTVENKPANNTPAQGNVDTPGSEDTMESRRSSMASTS 442
                                                                                                                                                                                                                                                                                                                                                              267 SASSSQAPQVSTSTDNTTSNASASNSSNSSNSSSSSSS----SSSSSSSSSNSNAGGNT 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503 DNGARL----LGNPSA------GIQSTYARLALSG-------GLRHDMG 534
                                                                                                           284 -----TKEAFQNPDNQKVNIDE---LGNAIPS-----GVLKDDVVANIEEQAKA 324
                                                                                                                                                                                                                                                                    181 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTST--SSLRSDPK-LWLALGT 237
                                 32 DIAKODATISSAQSAKAQAQAQVDSLQSKVDSLQQKQTSTKAQIAKIESERKALNAQIAT 91
                                                                                                                                                                                                                                                                                                                                                                                                    443 STFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPPRDTT
                                                                                   ----AASATETATRDQL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: St. Geme III, Joseph W.
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08/296,791
FILING DATE: 25.40G-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
                                                                                   238 VATGLIGLAATGIVQALALTPEPDSPTTTDPDA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 5, Application US/08296791
Patent No. 6245337
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   535 ---GLTGGSNSA--VNTSNNP 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         396 AAPGVIASSFSADFVGYANSP 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (415) 781-1989
(415) 398-3249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 1702 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                        323 NSGTSTGNTGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-296-791-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                     ð
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1320 -----INTGSATAITETAEKSDKPQTETAA 1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1242 QASPKQEQSETVQPQAVLESENVPTVNNA-----EEVQAQLQTQTSAT-VSTKQPAPE 1293
                                                                                                                                                                                1200 -AEP-------VTPVVSKNQTENTTDQPTEREKTAKVETEKTQEPPQVAS 1241
                                                                                                                                                                                                                                                                       ------DEIQEAPQMASE 1141
                                                                                                                     ------TVAENSKQES--- 1053
                                                                                                                                                                                                                                                                                                          182 LLEPKGTGESKGAGESKGVGELR-ESNSGAENTT-------ETQTSTSSL 225
                                                                                                                                                                                                                                                                                                                                                                                 226 RSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTK 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                          EAFQNPDNQKV-----NIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIE 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336 NNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTT 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           396 TITITITISARIV-ENKPANNIPA-----QGNVDIPGSEDIMESRRSSMASISSIFFD 447
                                                                                                                                                                                                                                  122 GQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRILE 181
                                                                                                                                                         62 RASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAV 121
                                                                                10 PNVNNSI-----PPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDN 61
1 4.4%; Score 125; 2.7. Similarity 17.8%; Pred. No. 0.044; Similarity 17.8%; Pred. No. 0.044; Indels 160; Mismatches 180; Indels 160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CVERRATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/106619
                                                                                                                                                                                                                                                                 1100 TQTTEIKETAKVEKEEKAKVEKEEKAKVEK-----
                                                                                                                       1017 PSNNEEIARVETPVPPAAPATPSETTE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,791
FILING DATE: 25-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1379 NSKRSKPNR 1387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              448 TSSIGGPCR 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
```

```
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-728-470-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                    17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1142 TSPKQAKPAPKEVSTDTKVEETQVQAQPQTQSTTVAAAEATSPNSKPAEETQPSEKTN-- 1199
                                                                                                                                                                                                                                                                                                                                                1054 -----KTVEKNEQDATETTAQNG-EVAEEAKPSVKANTQTN-----EVAQSGSETEE 1099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1100 TQTTEIKETAKVEKEEKAKVEKEEKAKVEK-----------DEIQEAPQMASE 1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1200 -AEP------VTPVVSKNQTENTTDQPTEREKTAKVETEKTQEPPQVAS 1241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      226 RSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTK 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              286 EAFQNPDNQKV------NIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIE 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    336 NNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTT 395
                                                                                                                                                                                                                                                                                                                                                                                              62 RASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAV 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 LLEPKGTGESKGAGESKGVGELR-ESNSGAENTT--------ETQTSTSTSL 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             396 TITITITISARIV-ENKPANNIPA-----QGNVDIPGSEDIMESRRSSMASTSSTFFD 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       122 GQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRILE 181
                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                       10 PNVNNSI-----PPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDN 61
                                                                                                                                                                                                                                                                                    Indels 160;
                                                                                                                                                                                                                                           Query Match
4.4%; Score 125; DB 5; Length 1702;
Best Local Similarity 17.8%; Pred. No. 0.044;
Matches 87; Conservative 62; Mismatches 180; Indels 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 10, Application US/08728470
Patent No. 5928651
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
         FP-59941/RFT
REFERENCE/DOCKET NUMBER: FP-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISICS:
LENGTH: 1702 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                        unknown
                                                                                                                                                      TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1379 NSKRSKPNR 1387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        448 TSSIGGPCR 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CITY: Arlington STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A. ZIP: 22202-0286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1294 NS-----
                                                                                                                                                                    ; TOPOLOGY:
PCT-US95-10661A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-728-470-10
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
1134 ISTKTGDIKGGIESTSGNVNITASGNTLKVSNITGQDVTVTADAGALTTTAGSTISATTG 1193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1309 TTQTGSSITSSNGQTTLTAKD-SSIAG------1335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  882 TINASGIQKTIINGNIINEKGDLN-----IKNIKADAEIQIGGNISQKEGNLTISSDK 934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        935 VNITNQITIKAGVEGGRSDSSEAENANLTIQTKELKLAGDLNISGFNKAEITAKNGSDLT 994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTVAT 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------PDNQKVNIDELGNAIP-SGVLKDDVVANIEEQAKAAGEEAKQQAIEN 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAQAQKKYDEQQAKRQEELKVSSGA-GYGLSGALILGGGIG---VAVTAALHRKNQPVEQ 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           487 ISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26 TDGAGG-----RQQLINSTGPLGSRALFTPVRNSMADS----GDNRASDVPGLPVNPMR 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1079 KEGITINATIGSVEVIA----QNGTIKGNITSQNVTVTATENLVITENAVINATSGTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 995 IGNASGGNAD-----AKKVTFD-KVKDSKISTDGHN-----VTLNSEV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         393 TITITITITISAR----TVENKPANNTPAQGNVDTPGSEDIMESRR------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --SSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 --GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKE-AFQN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1529;
                          SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 4.4%; Score 124.5; DB 2; Best Local Similarity 20.2%; Pred. No. 0.041; Matches 123; Conservative 76; Mismatches 233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             L-AASEITLNDGFE-VLHDHGPLDTLNRQIGSSVFRVE---
                                                                                                                                                                                                                                    APPLICATION UNBER: US 08/302,832
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION UNBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                     US/08/728,470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Berkstesser, Jerry W
REGISTRATION NUMBER: 22,651
REPERENCE/DOCKET NUMBER: 1038-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
INFORMATION FOR SEC ID NO: 10:
SEQUENCE CHARACTERISTICS:
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 1529 amino acids
                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
```

```
----AKKVTFD-KVKDSKISTDGHN--
                                                                                                                                                                                                                                                                                                                                                                                                                              1309 TTQTGSSITSSNCQTTLTAKD-SSIAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION DATA:
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1387 TN---ASGS 1392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            547 SNNPPAPGS 555
   995 IGNASGGNAD --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
US-08-617-697-10
                                                                                                                                                                                                                                                                                                                                                                                                  436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                       ŏλ
                                                                       qq
                                                                                                          ōλ
                                                                                                                                            qq
                                                                                                                                                                                δ
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                            ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
         -----TQEDGKHIA 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 177; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26 TDGAGG-----RGQLINSTGPLGSRALFTPVRNSMADS----GDNRASDVPGLPVNPMR 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1529;
                                                                                                                                                                              Sequence 10, Application US/08719641

Sequence 10, Application US/08719641

Batent No. 6218141

GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 6218141-Typeable Haemophilus
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                          ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
4.4%; Score 124.5; DB 4;
Best Local Similarity 20.2%; Pred. No. 0.041;
Matches 123; Conservative 76; Mismatches 233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             76 L-AASEITLNDGFE-VLHDHGPLDTLNRQIGSSVFRVE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US PCT/US93/02166 FILING DATE: 16-MAR-1993 PRIOR APPLICATION DATA: APPLICATION NUMBER: GB 9205704.1 FILING DATE: 16-MAR-1992 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Berkstresser, Jerry W
REGISTRATION NUMBER: 22.651
REFERENCE/DOCKET NUMBER: 1038-625
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFANE: (703) 415-0813
INFORMATION FOR SEQ ID NO: 10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 1529 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                         STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS: si
                                                                                             1387 TN---ASGS 1392
                                                                                                                                                                                                                                                                                                                                                                                        CITY: Arlington STATE: Virginia
                                                               547 SNNPPAPGS 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; TOPOLOGY:
US-08-719-641-10
                                                                                                                                                         RESULT 5
US-08-719-641-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
1336 INAANVTLNTTGTLTTTGDSKI--NATSGTLTINAKDA-----KLDGAASGDRTVVNA 1386
                                                                                                                                                                                                                                                                                                                                                                                                                  1250 TINGTNSVITSSQSGDIEGTISGNTVNVTASTGDL-TIGNSAKVEAKNGAATLTAESGKL 1308
                                             1032 -----RISNGSSNAGNDNSTG----LIISAKDVTVNNNVTSHKTI-----NISAAAGNVTT 1078
                                                                                                                                                                   1079 KEGTTINATTGSVEVTA----QNGTIKGNITSQNVTVTATENLVTTENAVINATSGTVN 1133
                                                                                                                                                                                                                                                                                               -------PDNQKVNIDELGNAIP-SGVLKDDVVANIEEQAKAAGEEAKQQAIEN 336
                                                                                                                                                                                                                                                                                                                                                                       337 NAQAQKKYDEQQAKRQEELKVSSGA-GYGLSGALILGGGIG---VAVTAALHRKNQPVEQ 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----SSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQ 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         487 ISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNT
181 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVAT
                                                                                                                        241 --GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKE-AFQN---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 10, Application US/08617697
Sequence 10, Application US/08617697
Batent No. 5977336
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                393 TTTTTTTTTTSAR-----TVENKPANNTPAQGNVDTPGSEDTMESRR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 05-0CT-1994
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: BETKSTESSER, JESTY W
REGISTRATION NUMBER: 22,651
RESPERDECE/DOCKET NUMBER: 22,651
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
```

```
unknown
                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                ;
US-08-409-995-5
                                                                                                                                                                                                                                                                                                                                                                              LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----KTSNGSSNAGNDNSTG----LTISAKDVTVNNNVTSHKTI----NISAAAGNVTT 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEGTTINATTGSVEVTA----QNGTIKGNITSQNVTVTATENLVTTENAVINATSGTVN 1203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1320 TINGTNSVTTSSQSGDIEGTISGNTVNVTASTGDL-TIGNSAKVEAKNGAATLTAESGKL 1378
                                                                                                                                                                                                                                                                                                                                                          1005 VNITNQITIKAGVEGGRSDSSEAENANLTIQTKELKLAGDLNISGFNKAEITAKNGSDLT 1064
                                                                                                                                                                                                                                                                                    952 TTNASGTQKTIINGNITNEKGDLN-----IKNIKADAEIQIGGNISQKEGNLTISSDK 1004
                                                                                                                                                                                                                                                                                                                         76 L-AASEITLNDGFE-VLHDHGPLDTLNRQIGSSVFRVE-------TQEDGKHIA 120
                                                                                                                                                                                                                                                                                                                                                                                           121 VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241 --GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKE-AFQN----- 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------PDNQKVNIDELGNAIP-SGVLKDDVVANIEEQAKAAGEEAKQQAIEN 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAQAQKKYDEQQAKRQEELKVSSGA-GYGLSGALILGGGIG---VAVTAALHRKNQPVEQ 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----SSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQ 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            487 ISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNT 546
                                                                                                                                                                                                                                                 TDGAGG-----RGQLINSTGPLGSRALFTPVRNSMADS----GDNRASDVPGLPVNPMR 75
                                                                                                                                                                                                                                                                                                                                                                                                                 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSTRSDPKLWLALGTVAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TTTTTTTTTTTSAR-----TVENKPANNTPAQGNVDTPGSEDTMESRR------
                                                                                                                                                                                                                Indels 177;
                                                                                                                                                                             Length 1600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Barenkamp, Stephen I.
APPLICANT: St. Geme III, Joseph W.
TITLE OF INVENTION: Haemophilus Adhesion Proteins
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
                                                                                                                                                                           Query Match 4.4%; Score 124.5; DB 2; Best Local Similarity 20.2%; Pred. No. 0.044; Matches 123; Conservative 76; Mismatches 233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1379 TTQTGSSITSSNGQTTLTAKD-SSIAG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 7
US-08-409-995-5
Sequence 5, Application US/08409995
Patent. No. 5646259
GENERAL INFORMATION:
BATCHAL INFORMATION:
BATCHAL INFORMATION:
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 amino acids
                                                                     TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       San Francisco
                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1457 TN---ASGS 1462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  547 SNNPPAPGS 555
                                                                                                        ; TOPOLOGY:
US-08-617-697-10
                                                                                                                                                                             Query Match
Best Local 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1264
                                                                                                                                                                                                                                                   56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   393
                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                           ογ
```

```
334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           171 DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----FATVASGINVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484 TTAAEADGGTLDGNASEQEVKAGDKVTFKAGKNLKVKQEGANFTYSLQDALT---GL--- 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194 PKVN------VTSTTDGLKFAKDAAGANGDTTVHLNGIGSTLTDTLVGSPATHIDGGD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         111 ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                441 ---LTVNDGKNANN-----PKGKVAD--VASTDEKKLVTAKGLVTALNSLSWT ,483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        379 VTAALHRKNOPVEQTITITITITITISARTVENKPA-NNTPAQGNVDTPGSEDTMESRRSS 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 PNVNNSIPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  305 DSKENGKRTEV--KIGAKTSVI------KEKDGKLFTG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.3%; Score 123; DB 1; Length 658;
22.4%; Pred. No. 0.015;
.ve 58; Mismatches 171; Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK--
                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/409,995
                                                                                                                                                                                                                                                                                                                                 REFERENCE/DOCKET NUMBER: A-61053/RET
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                      UMBER: US/08/409,995
24-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-685-467-5; Sequence 5, Application US/08685467; Patent No. 6060059
                                                                                                                                                                                             FILING DATE: 24-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 22.4%;
Matches 111; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    658 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 570 NANTISVTKDGISAGG 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             double
COMPUTER READABLE FORM:
```

```
379 VTAALHRKNOPVEOTTTTTTTTTTTSARTVENKPA-NNTPAQGNVDTPGSEDTMESRRSS 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.3%; Score 123; DB 4; Length 658; 22.4%; Pred. No. 0.015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,942
FILING DATE: 29-DEC-1997
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: St. Geme, Joseph
APPLICANT: Barenkamp, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DAV TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             58; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
PRIOR APPLICATION NUMBER: PCT/US96/4031
APPLICATION NUMBER: PCT/US96/4031
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vance, Dolly A.
REGISTRATION NUMBER: 39,054
                                                                                                                                                                                                                                  RESULT 9
US-08-913-942-5
; Sequence 5, Application US/08913942
; Patent No. 6200578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          658 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 22.4 Matches 111; Conservative
                                                                                                                                                    570 NANTISVIKDGISAGG 585
                                                                                                         438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             , MOLECULE TYPE: protein US-08-913-942-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111
                                                            538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                            g
                                                                                                              δ
                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 194 PKVN------VTSTTDGLKFAKDAAGANGDTTVHLNGIGSTLTDTLVGSPATHIDGGD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 111 ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         171 DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :| : | : | : | 385 ---KANK---ETNKVDGANATEDADEGKGLVTAKDVIDAVNKTGWRIKTTDANGQNGD-- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     387 ----FATVASGINVIFASGNGTIATVINGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK------ 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        324 AAGEEAKQQAIENNAQAQ -- KKYDEQQAKRQEELKV - SSGAG -- YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10 PNVNNSIPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.3%; Score 123; DB 3; Length 658; 22.4%; Pred. No. 0.015;
                                                                   TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS NUMBER OF SEQUENCES: 6 CORRESPONDENCE ADDRESS: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Floppy disk
COMPUTER: TBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,467
FILING DATE: 22-JUL-1996
CLASSIFICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
ATTONREY/AGENT INFORMATION:
NAME: S11Va, RODIN M.
REGISTRATION NUMBER: 38,304
                 APPLICANT: St. Geme III, Joseph W.
APPLICANT: Barenkamp, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            58; Mismatches 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (415) 781-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 658 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 4.35
Best Local Similarity 22.45
Matches 111; Conservative
                                                                                                                                                                                          CITY: San Francisco
STATE: California
COUNTRY: United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , MOLECULE TYPE: protein US-08-685-467-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid STRANDEDNESS: un
GENERAL INFORMATION:
```

q

셤 õ g ö g ö g ò 음 à g

```
25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  246 QSTHYTRAASIKDVLNAGWNIKGVK-AGSTIGQSENVDFVHTYDTVEFLSADTETTTVTV 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  194 PKVN------VTSTTDGLKFAKDAAGANGDTTVHLNGIGSTLTDTLVGSPATHIDGGD 245
------AGAN 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 PNVNNSIPPAPPLPSQTDG-----AGGRQQLINSTGPLGSRALFTPVRN--SMADSGD
```

```
538 ------AGANGAKTEINKDGLTITPANG-----AGAN 569
                                                                                                                          246 QSTHYTRAASIKDVLNAGWNIKGVK-AGSTIGQSENVDFVHTYDIVEFLSADTETITVTV 304
                                                                                                                                                                                       111 ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                 :| : | : | : | : | 335 ---KANK---ETNKVDGANATEDADEGKGLVTAKDVIDAVNKTGWRIKTTDANGQNGD-- 386
                                                                                                                                                                                                                                                                                                                                                                            ----TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                          ----FATVASGTNVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   194 PKVN------VTSTTDGLKFAKDAAGANGDTTVHLNGIGSTLTDTLVGSPATHIDGGD 245
                                                                                                                                                                                                                                                                                 171 DITEARORILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                379 VTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPA-NNTPAGGNVDTPGSEDTMESRRSS 437
    10 PNVNNSIPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD
                                                                                               ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 484 TTAAEADGGTLDGNASEQEVKAGDKVTFKAGKNLKVKQEGANFTYSLQDALT---GL---
                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Barenkamp, Stephen J.
TITLE OF INFORMATION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 94111-4187
COMPUTER READABLE PORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEPA: (415) 398-3249
TELEX: 910 277299
                                                                                                                                                                                                                                                                                                                                                                            231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/685,467 FILING DATE: 22-JUL-1996 CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.08-685-467-2
Sequence 2, Application US/08685467
Patent No. 6060059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38,304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        570 NANTISVTKDGISAGG 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: United States ZIP: 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Silva, Robin M. REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: San Francisco
STATE: California
                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                            387
                                                                                                                                   g
                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                               ò
                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25;
                                                                                                                                                                                                                                                                171 DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPK 230
                                                                                                                                                                                       387 ----FATVASGTNVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                             324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                         484 TTAAEADGGTLDGNASEQEVKAGDKVTFKAGKNLKVKQEGANFTYSLQDALT---GL--- 537
                                                                                                                                                                                                                                                                                                                                                                                                                          379 VTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPA-NNTPAQGNVDTPGSEDTMESRRSS 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                     ---KANK----ETNKVDGANATEDADEGKGLVTAKDVIDAVNKTGWRIKTTDANGQNGD--
                                                                                                                                        231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 4.3%; Score 123; DB 1; Length 1098; Best Local Similarity 22.4%; Pred. No. 0.033; Matches 111; Conservative 58; Mismatches 171; Indels 156;
                                                                                                                                                                                                                                   280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert STREET: Four Embarcadero Center, Suite 3400 CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARES PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/409,995
FILING DATE: 24-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: S11va, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/POCKET NUMBER: A-61053/RFT
TELEPHONINICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 781-1989
TELEFX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
       -KEKDGKLFTG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Barenkamp, Stephen I.
APPLICANT: Bt. Geme III, Joseph W.
TITLE OF INVENTION: Haemophilus Adhesion Proteins
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08409995
Patent No. 5646259
305 DSKENGKRTEV--KIGAKTSVI--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1098 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   570 NANTISVTKDGISAGG 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER: IBM PC COPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: unknown US-08-409-995-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: do
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SULT 10
-08-409-995-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                            335
                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                          g
                                                                                                                                        ð
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
```

```
STREET: Four Embarcaderc
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          570 NANTISVTKDGISAGG 585
                                                                                                                                                          Matches 111; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA
                                                                                                                             Similarity
US-09-377-155-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-913-942-2
                                                                                            Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                       25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK------ 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        379 VIAALHRKNQPVEQTITITITITITISÅRTVENKPA-NNIPAQGNVDTPGSEDIMESRRSS 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----FATVASGINVTFASGNGTTATVINGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      484 TTAAEADGGTLDGNASEQEVKAGDKVTFKAGKNLKVKQEGANFTYSLQDALT---GL--- 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   111 ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        194 PKVN-----VTSTTDGLKFAKDAAGANGDTTVHLNGIGSTLTDTLVGSPATHIDGGD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 246 OSTHYTRAASIKDVLNAGWNIKGVK-AGSTTGQSENVDFVHTYDTVEFLSADTETTTVTV 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 PNVNNSIPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                         Query Match
4.3%; Score 123; DB 3; Length 1098;
Best Local Similarity 22.4%; Pred. No. 0.033;
Matches 111; Conservative 58; Mismatches 171; Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: PEAK, Ian Richard Anselm APPLICANT: JENNINGS, Michael Paul APPLICANT: JENNINGS, Michael Paul APPLICANT: MONON, E. Richard Paul CANTON ON THE STREET OF STRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 32, Application US/09377155 Patent No. 6197312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Haemophilus influenzae
                                                                                                                                                                                           ; MOLECULE TYPE: DNA (genomic) US-08-685-467-2
             INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1098 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                438 MASTSSTFFDTSSIGG 453
                                                                                                                                             unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQ ID NOS:
SOFTWARE: Patentin Ve
                                                                                                                                                                        unknown
                                                                                                         TYPE: amino acid
STRANDEDNESS: un
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 32
LENGTH: 1098
                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-377-155-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

```
25;
                                                                                                                                                                                                                                                                                                                                                                                              171 DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----FATVASGTNVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  484 TTAAEADGGTLDGNASEQEVRAGDKVTFKAGKNLKVKQEGANFTYSLQDALF---GL--- 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----TSITLGTGNNGAKTEINKDGLTITPANG-------AGAN 569
                                                                                                                                                                                                                                                                                             111 ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVAS 170
                                                                                                                                  61 ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                           246 QSTHYTRAASIKDVLNAGWNIKGVK-AGSTTGQSENVDFVHTYDTVEFLSADTETTTVTV 304
                                               Gaps
                                                                                           10 PNVNNSIPPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                                                             379 VTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPA-NNTPAQGNVDTPGSEDTMESRRSS
                                                                                                                                                                                                                                                                                                                          305 DSKENGKRTEV--KIGARTSVI------KEKDGKLFTG------
                                               58; Mismatches 171; Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: St. Geme, Joseph
APPLICANT: St. Geme, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
Score 123; DB 4;
Pred. No. 0.033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,942
FILLING DATE: 29-DEC-1997
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/08913942 Patent No. 6200578 GENERAL INFORMATION:
  4.38;
```

σ

```
: | | : | | : | 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----QA---EPAPEEVPTDTNAEEAQALQOTQPTTVAAAETTSP 1253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1310 SPKQEQPAAKPQAQTKPQAEP------ARENVLTTKNVGEPQPQAQPQTQ 1353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1254 NSKPAEETQQP-SEKTNAEPVTPVVSENTATQP--TETEETAKVE-KEKTQEVPQVASQE 1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------VTVSEN 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 172 ITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKL 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         232 WLALGTVATGLIGLAATGIVQALALIPEPDS-PTTTDPDAA------ASATETATR 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               281 DQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        341 QKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKN----QPVEQTTTT 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 397 TTTTTTTSARTVENKPA-----NNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFD 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       57 DSGDNRASDVPGLPVNPMRLA----ASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVE 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112 TQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASD 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17 PPAP--------PLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMA 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          52; Mismatches 174; Indels 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.3%; Score 123; DB 4; Length 1848; 21.3%; Pred. No. 0.076;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1147 TOSEGK------TEETOTAETKSEPTES-----
                                                                                                                                                                                   COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DAIA:
FPILING DATE: US/08/296,791
FLILIG DATE: 25-AUG-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
     : 4 Embarcadero Center, Suite 3400
San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1848 amino acids
                                                                                  COUNTRY: United States ZIP: 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 103; Conservative
                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                            California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1413 TSS 1415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
PCT-US95-10661A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          448 TSS 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-296-791-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                         STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----NRASDV-----PGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---KANK----ETNKVDGANATEDADEGKGLVTAKDVIDAVNKTGWRIKTTDANGQNGD-- 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         231 LWLALGTVATGLIGLAATGIVQALALTPEPDSPTT------TDPDAAASATETAT 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----FATVASGINVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADT-TA- 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  324 AAGEEAKQQAIENNAQAQ--KKYDEQQAKRQEELKV-SSGAG--YGLSGALILGGGIGVA 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----TSITLGTGNNGAKTEINKDGLTITPANG------AGAN 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194 PKVN------VTSTIDGLKFAKDAAGANGDITVHLNGIGSTLIDTLVGSPATHIDGGD 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      246 QSTHYTRAASIKDVLNAGWNIKGVK-AGSTTGQSENVDFVHTYDTVEFLSADTETTTVTV 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGAGHAMVTVAS 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPK 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            280 RDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQ----AK------ 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           379 VIAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPA-NNTPAQGNVDTPGSEDTMESRRSS 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 4.3%; Score 123; DB 4; Length 1098;
Best Local Similarity 22.4%; Pred. No. 0.033;
Matches 111; Conservative 58; Mismatches 171; Indels 156; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10 PNVNNSIPPAPPLPSQTDG-----AGGRGQLINSTGPLGSRALFTPVRN--SMADSGD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         305 DSKENGKRIEV---KIGAKTSVI-----KEKDGKLFTG-----KEKDGKLFTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 6, Application US/08296791
Patent No. 6245337
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetrati
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
                                                                                  NAME: Vance, Dolly A.
REGISTRATION NUMBER: 39,054
REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DAV
TELECOMNUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
APPLICATION NUMBER: PCT/US96/4031 FILING DATE: 22-MAR-1996 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                TELEPHONE: (415) 781-1989
TELEEX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
LENGTH: 1098 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     570 NANTISVTKDGISAGG 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 438 MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-913-942-2
                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-296-791-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               538
```

ò g ò ò

ç

```
23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : | | : | | : | | : | | 1054 NSKPAEETQQP-SEKTNAEDVTPVVSENTATQP--TETEETAKVE-KEKTQEVPQVASQE 1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : | | : | | : | 1093 VPSETTEKVAENPPQENETVAKNEQEATEPTPQNG-EVAKEDQPTVEANTQTNEA---- 1146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------OPEKTVSQ--STEDKVVVE-KEEKAKVE-TEETQKAPQVTS-KEPPK- 1216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1038 PPAPATESAIASEOPETRPAETAOPAMEETNTANS----TETAPKSDTATOTENPNSES 1092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1147 TQSEGK------VTVSEN 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     281 DQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQA 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          341 QKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKN----QPVEQTITT 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              232 WLALGTVATGLIGLAATGIVQALALTPEPDS-PTTTDPDAA------ASATETATR 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        57 DSGDNRASDVPGLPVNPMRLA----ASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVE 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               112 TQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASD 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         172 ITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKL 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
4.3%; Score 123; DB 5; Length 1848;
Best Local Similarity 21.3%; Pred. No. 0.076;
Matches 103; Conservative 52; Mismatches 174; Indels 154; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PPAP------PLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMA 56
                       APPLICANT: Washington University, et al.

TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
ADDRESSE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco
                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REFERENCE/DOCKET NUMBER: 31,801
REGISTRATION NUMBER: FP-59941/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SED ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                            NUMBER: PCT/US95/10661A
16-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,791
FILING DATE: 25-AUG-1994
Sequence 6, Application PC/TUS9510661A GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 1848 amino acids
TYPE: amino acid
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                  CURRENT APPLICATION DATA:
APPLICATION NUMBER: PC1
FILING DATE: 16-AUG-199
CLASSIFICATION:
                                                                                                                                                                                 CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US95-10661A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

THIS PACK BLANK USANO,

~

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

September 27, 2001, 14:21:42; search time 25.62 Seconds (without alignments) 1662.044 Million cell updates/sec Run on:

US-09-189-415A-4 2851 1 MPIGNLGHNPNVNNSIPPAP......SNSAVNTSNNPPAPGSHRFV 559 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

219241 segs, 76174552 residues

Searched:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_68:* Database :

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description				hypothetical prote	e memb	related to C2H2 zi	probable serine/th	probable tail fibe	hypothetical prote		ascites sialoglyco	mucin homolog - bo	lactocepin (EC 3.4	maternal effect pr				c	probable adhesin 2	probable adhesin Z	hemagglutinin/hemo	hypothetical prote	hypothetical prote	lary	secreted 45K prote	pothe	hypothetical prote	О	hypothetical prote
QI OI	E86045	2569	443	T13389	C85693	T51024	T18611	G64887	T34513	T33369	A53577	A36054	A32634	A40315	T24583	A83412	T02761	C82199	н85611	B85663	F81045	T21460	E71497	T42233	JN0097	H86561	B72061	JQ0188	T32271
DB	7	~	~							~	~	~	~	-	~	7	~	7	~	7	7	~	~	~	~	~	~	7	~
Length	558	1229	2232	1291	973	770	1192	1122	3507	1275	1630	563	1962	1026	1829	2468	837	4558	1005	1005	2514	196	439	1589	461	755	755	1258	1459
% Query Match	92.1		•	•	•	5.5		•	•	•		•			4.9	•		4.8	4.8	•		4.7	•	4.7	4.6	•		4.6	4.6
Score	2624.5	_			152.5	148.5	*	146.5	4		143.5	142	141.5	140.5	140.5	139	137	137	136	136	$\overline{}$		133	m				132.5	
Result No.		7	m	4	2	9	7.	80	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

qq à d

hypothetical prote	cell surface glyco	probable invasin Z	probable adenylate	ice nucleation act	hypothetical prote	serine-rich protei	extracellular matr	surface protein XF	tyrosine kinase su	flagellar hook-ass	SPR-1 protein - hu	polymorphic membra	CREB-binding prote	hypothetical prote	conserved hypothet	
T31597	JC2300	E85822	D71485	JC2143	D86731	A44067	T31110	D82671	A46299	G82111	S26638	D81675	T13828	T06029	G81655	
7	~	7	~	~	ď	~	~	~	7	~	~	~	7	7	7	
069	1083	2660	829	1034	1063	1077	2055	2059	2411	999	784	1460	3190	532	436	
4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	
131	131	131	130.5	130.5	130	130	130	129.5	129.5	129	129	129	129	128.5	127.5	
30	31	32	33	34	35	36	37	38	39	. 40	. 41	4.5	4.3	44	45	

ALIGNMENTS

20

```
A; Molecule type: DNA
A; Residues: 1-2232 <GEI>
                                                                                                                          C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Realdues: 1-1229 GFUL>
A;Costatues: 1-1229 GFUL>
A;Costatues: EMBL:U67956; PIDN:AAB07691.1; GSPDB:GN00028; CESP:F16F9.2
A;Experimental source: strain Bristol N2; clone F16F9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
T34434
Hypothetical protein K06A9.la - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C;Accession: T34434
A;Reference number: 221525
A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hypothetical protein F16F9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 *sequence_revision 15-Oct-1999 *text_change 18-Feb-2000
C;Accession: T25697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQ 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LTKEAF-----QNPDNQKVNIDELGNAIPSGVLKDDV-VANIEEQAKAAGEEAKQQ 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            705 ITQTSVSVVESSTPRQLPERWKAIVNKFKHNLE--VLKEKKRLLKEKESISTTGSDSS-- 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          333 AIENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQ 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TVASDITEARQRILELLEPKGTGESKGAGESKGVGELR----ESNSGAENTTETQTSTST 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SVSTEKTSTTKKASTTEEPTTTTESSTTGKATTPELSTTSEETTTTELKITTEG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----VOTTPITTEKSTTOEE 797
                                                                                                                                                      NVDTPGSEDTMESRRSSWASTSSTFFDTSSIG --- GPCRIRMLMLKHRCMIRRCRLLILL 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    393 TITITITITITASARTVENKPANNIPAQGNVDIPGSEDIMESRRSSMASTSSTFFDISS 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----TTETTTSEPSTTESTTVDTSS 844
                                                                    A;Map position: X
A;Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R;Fulton, B. submitted to the EMBL Data Library, August 1996 A;Description: The sequence of C. elegans cosmid F16F9. A;Reference number: 220071 A;Reference number: 725697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44; Mismatches 122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.0%; Score 170; DB 2; 22.8%; Pred. No. 0.025;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                761 --ETTTVVAENIDEVTTTEKEKV----
                                                                                                                                                                                                                                                                                                               GGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 22.8%
Matches 68; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :Gene: CESP:F16F9.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              657
                                                                                                                                                                                                                                                                                            538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                            ò
                                                                                                          g
                                                                                                                                                                        à
```

```
hypothetical protein 115C2.10 - fruit fly (Drosophila melanogaster)
C; Species: Drosophila melanogaster
C; Species: Janua-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C; Accession: Ti3389
R; Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D. submitted to the EMBL Data Library, May 1999
A; Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A; Reference number: Z17665
A; Cacession: Ti3389
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; canslated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-1291
A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacession: A; Cacess
                                                                                                                                                                                                                                       2049/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Residues: 1-1291 <CAT>
A; Cross-references: EMBL:AL031581; NID:e1320978; PID:e1426292; PIDN:CAA20894.2
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20;
A;Cross-references: EMBL:U80846; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K06A9.1a
A;Experimental source: strain Bristol N2; clone K06A9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1519 TGTMSSTSSGTVGSTISES-STTASASSQTGSTVTMGSSSTSGV-----STSSASSTQP 1571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1343 SPSSISPVPISSPIPSTIFASSTSGSTISDVSSVSTTSL-APLSSSLPSTVPSSTQSFSS 1401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            217 QTSTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATE 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      277 TATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIEN 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                337 NAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTT 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     397 ITTITISARIVENKPANNIPAQGNVDIPGSEDIMESRRSSMASISSIFFDISSIGGPCR 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             110 VETQEDGKHIAVG--QRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHA--- 164
                                                                                                                                                                                                                             A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56 ADSGDNRASDVP-----GLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFR 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----MVTVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1616 TISESSTTASA----SSQTGSTVTMGSSSTSGVSTSSA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               457 IRMIMLKHRCMIRRCRILILIRLFRIWGIQIS------VVYSTIQHPPRDTTDNGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1650 -----STQPQMSTSQGSS---AGSTVASSTTGLVSTSTV-----PSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1402 TSEGSSKASSSPVPSQTSSTPTNPTGSTESSTLLSSTISGSTQH---TTMSKASSGSTSP
                                                                                                                                                                                                                                                                                                                                                                                         Length 2232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1769 GTMSSTSSGTVGSTMSQSSTAASTTSHTGSTVTLGSSS---TSSN 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              507 RLLGNPSAG-IQSTYARLALSGGLRHDMGG-LTGGSNSAVNTSNN 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9 NPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSM--
                                                                                                                                                                                                                                                                                                                                                                                              5.8%; Score 165.5; DB 2;
19.3%; Pred. No. 0.1;
ative 81; Mismatches 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Cross-references: FlyBase:FBgn0020381
A;Map position: X
                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.8%
Best Local Similarity 19.3%
Matches 113; Conservative
                                                                                                                                                                     A; Gene: CESP: K06A9.1a
```

A;Introns: 238/3; 1225/1 A;Note: EG:115C2.10	182 LLEPKGIGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDFKLWLALGTVATG
Query Match 5.7%; Score 162.5; DB 2; Length 1291; Best Local Similarity 20.6%; Pred. No. 0.075; Matches 99; Conservative 61; Mismatches 189; Indels 131; Gaps 18;	157SATSAGGAASSAGSASSAGTASTKATEASKSAAABESSKSAAAT 242 LIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFONPDNGKVNIDEL
Qy 5 NLGHNPNVNNSIPPRAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDN 61	
QY 62 RASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQ 102	PAQ-
QY 103 IGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAG 162	420GNVDTPGSEDTMESRRSSMASTSSTFFDTSS 450
Qy 163 HAMVIVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTST 222	DD 360 AKTSETNAKASETSAESSKTAAASSASSAAS 393 RESULT 6
QY 223 SSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQ 282	T51024
QY 283 LTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQA 333 L : : : : : : : </td <td>C;Date: 21-Uul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000 C;Accession: T51024 C;Accession: T51024 R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu submitted to the Protein Sequence Database, July 2000</td>	C;Date: 21-Uul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000 C;Accession: T51024 C;Accession: T51024 R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu submitted to the Protein Sequence Database, July 2000
QY 334 IENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQT 393 1	A; Reference number: 225286 A; Accession: F1024 A; Status: preliminary A; Molecule type: DNA
Qy 394 TTTTTTTTTSARTVENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSS 450	A;Residues: 1-770 <sch> A;Cross-references: EMBL:AL389901; G\$PDB:GN00116; NCSP:B7F21.50 A;Experimental source: BAC clone B7F21; strain OR74A CGenetics: ACC Constant Source: BAC clone B7F21; strain OR74A</sch>
	A; Map position: 6 A; Introns: 117/1
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strai C;Species: Escherichia coli (strai C;Species: Escherichia coli (strai C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001 C;Accession: C85633	Query Match 5.2%; Score 148.5; DB 2; Length 770; Best Local Similarity 19.9%; Pred. No. 0.25; Matches 109; Conservative 62: Mismatches 151; Indels 227; Gaps 25;
R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew iller, L.; Grotheck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001. A; Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.	-DCAGGRGQLINSTGPLGSRALF 48 1
	Qy 49 TPVRNSMADSGDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSS 106
A; Kesloues: 1-9/3 <stoa A;Cross-references: GB:AE005174; NID:g12514847; PIDN:AAG56007.1; GSPDB:GN00145; UWGP:219 A;Experimental source: strain 0157:H7, substrain EDL933 C;Genetics:</stoa 	OY 107 VFRVETQEDGKHIDQ 136 i
A;Gene: Z1918	OY 137 EYARLOSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRILE 181
Questy Marcin Best Local Similarity 22.1%; Pred. No. 0.2; Matches 87; Conservative 55; Mismatches 169; Indels 83; Gaps 14;	182 LLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGT
79 SEITLNDGFEVLHDHGPLDTLNRQIGSSVF	443DSQSTGLGVHYSSTTRNQHQHHWRNQSNASELSADGESITH
Db 61 SVILLVEGFPFSHAGTITVYEDSQP-GTLNDFLGAMTEDDVRPEALRRFELMVEE 114 Qy 124RNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRILE 181	Qy 238 -VATGLIGLAATGIVQALALIPEPDSPITIDPD 259
Db 115 VARNASAVAQNTAAAKKSASDASTSAREAATHATDAADSARA 156	QY 270 AAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVA 316

09 146 PEGKDKEVFTGGGGGGHAMVTASDITEARQRILELLEPKGTGESKGAGESKGVGELRE 205 601GTRHGGVQMRAQPT-SRQATISLLQPPSYRPSSNTTQIAQIPPLEN 645 602 5NSGAENTTETGTSTSSLRSDPKLMIALGTVATGLIG	RESULT 8 G6487 G64887 G64887 G64887 G64887 G64887 G75becies: Escherichia coli C;Species: T09189 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A;Rosce, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997 A;Title: The complete genome sequence of Escherichia coli K-12. A;Reference number: A64720; MUID:97426617 A;Reference number: A64720; MUID:97426617 A;Residues: 1-1122 <blat> A;Roccession: G64887 A;Residues: 1-1122 <blat> A;Residues: 1-1122 <blat> A;Residues: 1-1122 <blat> A;Cross-references: GB:AE000234; GB:U00096; NID:91787633; PIDN:AAC74454.1; PID:917876 A;Experimental source: strain K-12, substrain M61655 A;Experimental source: strain K-12, substrain M61656 A;Experimental source: strain K-12, substrain M61667 /blat></blat></blat></blat>	A; Reference number: 216603; MUID: 97251357 A; Reference number: 216603; MUID: 97251357 A; Accession: T00189 A; Accession: T0189 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: DNA A; Residues: 3-1122 <alb> A; Residues: 3-1122 <alb< 3-1122="" 3-1222="" <a<="" <alb<="" a;="" residues:="" th=""></alb<></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb></alb>
: :	A. Status: translated from GB/EMBL/DDBJ A. Molecule type: DNA A. Molecule type: DNA A. Molecule type: DNA A. Residues: 1-487,536-1192 <will> A. Cross references: EMBL:Z81027; PIDN:CAB54178.1; GSPDB:GN00023; CESP:H39E23.1b A. Experimental source: clone AH10 B. Status: translated from GB/EMBL/DDBJ A. Reference number: Z19696 A. Reference number: Z19696 A. Status: translated from GB/EMBL/DDBJ A. Molecule type: DNA A. Molecule type: DNA A. Molecule type: DNA A. Residues: 1-1192 <wil3> A. Cross references: EMBL:Z96102; PIDN:CAB54263.1; GSPDB:GN00023; CESP:H39E23.1a A. Reperimental source: clone H39E23 A. Accession: T23143 A. Accession: T23143 A. Residues: 1-487,536-1192 <wil4> A. Residues: CESP:H39E23.1b A. SEXPERIMENTAL SOURCE: Clone H39E23.1b A. SEXPERIMENTAL SOURCE: Clone H39E23.1b A. SEXPERIMENTAL SOURCE: CLONE H39E23.1b A. Gross references: EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN00023; CESP:H39E23.1b A. SEXPERIMENTAL SOURCE: CLONE H39E23.1b A. Gross references: EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN00023; CESP:H39E23.1b A. Gross references: CESP:H39E23.1b A. Gross references: CESP:H39E23.1b</wil4></wil4></wil4></wil4></wil4></wil3></will>	A; Map position: 3 A; Map position: 3 A; Map position: 3 A; Map position: 10/3; 33/3; 67/3; 139/2; 183/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992 C; Networds: alternative splicing; ATP; phosphotransferase; serine/threonine-specific profess E; 1-192/Product: probable serine/threonine-specific protein kinase, long splice form #sprin-487,536-1192/Product: probable serine/threonine-specific protein kinase, short splice form #sprin-487, product splice form #sprin-487, product splice form #sprin-487, probable serine/threonine-specific protein kinase, short splice form #sprin-487, product

us-09-189-415a-4.rpr

Db 913 SDTTGKESSEETTSRKPIEGSDSLTEGSGGEWFETGSKGHFESGSKVSYTSGKGPTQSG 972 Qy 368 ALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSARTVENKPANNTPAGG- 420	RESULT 10 T33369 hypothetical protein H02F09.3 - Caenorhabditis elegans C.Species: Caenorhabditis elegans C.Date: 29-0ct-1999 #sequence_revision 29-0ct-1999 #text_change 17-Mar-2000 C;Accession: T33369 R;Geisel, C.; Harmon, G. Submitted to the EMBL Data Library, July 1998 A;Bescription: The sequence of C. elegans cosmid H02F09.	A; Reference number: 221330 A; Accession: T33369 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: DNA A; Residues: 1-1275 <gei>A; CEI>A; CEI A; Experimental source: strain Bristol N2; clone H02F09 C; Genetics: A; Gene: CESP: H02F09.3 A; Map position: X A; Introns: 42/1; 78/3; 106/3; 135/2; 161/2; 224/1 C; Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosida</gei>	Query Match Best Local Similarity 18.6%; Score 143.5; DB 2; Length 1275; Best Local Similarity 18.6%; Pred. No. 0.99; II indels 137; Gaps 15; Matches 95; Conservative 66; Mismatches 214; Indels 137; Gaps 15; QY 26 TDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNRASDVPGLPVNPM 74	OY 75 RLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120	QY 195 GESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQAL 254
	Qy 357 VSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPANNT 416 1	2K783.1 - Caenorhabditis elegans sittis elegans *sequence_revision 29-Oct-1999 #text_change 29-Oct-1999 b. M. L Data Library, August 1994 equence of C. elegans cosmid 2K783. 221536 ; translated from GB/EMBL/DDBJ	A; Residues: 1-3507 CRAV> A; Cross-references: EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1 A; Experimental source: strain Bristol N2; clone ZK783 A; Genetics: A; Markap position: 3 A; Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1; 1409/2; 3504/1	Query Match 5.1%; Score 144.5; DB 2; Length 3507; Best Local Similarity 19.7%; Pred. No. 3.4; Matches 105; Conservative 61; Mismatches 179; Indels 189; Ga 27 DGAGGRGOLINSTGPLGSRALFTPVRNSNADSGDNRASDVGLPVNPMRLAASEITL 11:	Db 787 SDGEESKGAGGAGHAMYTVASDITEARQRILELLEPKGTGESKGAGESKG 199 144 IDPEGKDKFVFTGGRGAGHAMYTVASDITEARQRILELLEPKGTGESKGAGESKG 199 15

ö

```
Jactococcus lactis subsp. cremoris lactococcus lactis subsp. cremoris lactococcus lactis subsp. cremoris N.Alternate names: cell envelope-associated serine proteinase prtP.
C;Species: Lactococcus lactis subsp. cremoris
C;Species: Lactococcus lactis subsp. cremoris
C;Bate: 21-May-1990 #text_change 21-Jul-2000
C;Accession: A32634
R;Vos. P.; Simons, G.; Siezen, R.J.; de Vos. W.M.
J; Biol. Chem. 264, 13579-13585, 1989
A;Title: Primary structure and organization of the gene for a procaryotic, cell envel A;Title: Primary structure and organization of the gene for a procaryotic, cell envel A;Title: Primary structure and organization of the gene for a procaryotic, cell envel A;Title: Primary structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, cell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, cell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, cell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, cell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic, sell envel A;Title: Primary Structure and organization of the gene for a procaryotic organization of the gene for a procaryotic organization organization of the gene for a procaryotic organization organization of the gene for a procaryotic organization 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Table 1999
C; Species: Bos primigenius taurus (cattle)
C; Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 05-Nov-1999
C; Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 05-Nov-1999
C; Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 05-Nov-1999
C; Accession: A36054
A; Proc. Natl. Acad. Sci. U.S.A. 87, 6798-6802, 1990
A; Title: Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein
A; Reference number: A36054, MUID: 90370871
A; Reference number: A36054
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-563 Asha>
A; Cross-references: GB:M36192; NID:g163400; PIDN:AAA30657.1; PID:g163401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13;
1187 QSYLTESMGASSTSETSLLTEATTETSLCLFTWTHCDRDLLSWTSSSGLTTKTDNDRSTA 1246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    374 GIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSARTVENKPANNTPAQGNVDTPGSEDTMES 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  195 GESKGYGELRESNSGA-ENTTETQTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQA 253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              149 TIIAPGSSNTKATIPIEVRTTIEVRIAIEIT-----TSRHSSDATGSGIQTGI---- 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 314 VVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGG 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    245 SSNTGTSTGVGR-----QTSTAVVSGRVTGV-SESSSPGTSKEASETTTGPGISTTGST 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           135 DQEYARLQSIDPECKDKFVFTGGRGAGHAMVTVASDITEARQRILELLEPKGTGESKGA 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88 EVLHDHGPLDTLNRQIGSSVFRVET------QEDGKHIAVGQRNGVETSVVLS 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :| :| | | | :: | | :: | | :: | | :: | | | :: | | :: | | | :: | :: | | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ------GTTK-VIPGTTVAPGSSNTESTTSLGESRTRIGRI-----TGATTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               254 LALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 142; DB 2; Length 563;
Pred. No. 0.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               48; Mismatches 169;
                                                                                                                                                                                                 1247 LSATSLTLPAPSTSTASRSTVPPAP 1271
                                                                                                             536 -----LTGGSNSAVNTSNNPPAP '553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 20.38
Matches 75; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          :|: :|||
SKSNRITISS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RRSSMASTSS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                         mucin homolog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bifun
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C.Accession: A55377
R.Wu, K.; Fregten, N.; Carraway, K.L.
J. Biol. Chem. 269, 11950-11955, 1994
A.;Title: Molecular cloning and sequencing of the mucin subunit of a heterodimeric, A.;Reference number: A53577; MUID:94216302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1139 IDSTSGNAGHTLLTGSHSVITSRVASTTLGRL-----STVAHRQSTQRSSTHS 1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              991 TOKKS--TVSTLVTS---TQELTSSQSQRTGSMGTSSK---------PQATTP 1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1030 TEVTISTLSSFS-----RGSLFSARNCCLQTKKPPLPAVVCLPDPSSVPSLMHSSKPQ 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TSSIGGPCRIRMLMAKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQH---PPRDTTDN 504
                                                                                                                                                                                                                                                                                                                                                                              ascites sialoglycoprotein 1 - rat (fragments)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 07-Feb-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QTTTTTTTTTTTSA----RTVENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFD 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 816 LSSSPSGSTPVQTRSVTSSSDERTNPTSSGVSNTSPATTEVLTPTSSPESTPGNTAPRTT 875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ETQISISISSLRSDPKLWLALG-----TVATGLIGLAATGIVQALALTPEPDSPTTT 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ETSTITITKVLMTSLOOKLPIGSILGTSTPTEVITILSASSSDOVOV----ETTSOTIL 930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SATE----TATRDQ----LTKEAFQNPD 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     931 SPDATITSHAPRESSSPPSTSVILTIMASTEGTSGDTGHTTAVTDQGSTPATTEISVTPS 990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nokvnidelgnaipsgvlkddvvanieeqakaageeakooaiennaqaqkkydeqqakro 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353 EELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPV---------E 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S-------PVSTFSTSTQEITTLS--QSQHTGGMKTTRNPQTTGTTEVTTTLSA 695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              595 HAPRESSS----PPSTSDILTTMASTEGTSGDTGHTMAVTTQGSTPATTEISVTPSTQKM 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADSGDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQED 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GKHIAVGQRNGVETSV---VLSDQEYARLQSIDPEGKDKFVFT----GGRGGAGHA---- 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 267; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HNPNVNNSIPPAPPLPSQ-----TDGAGGRGQLINSTG--PLGSRALFTPVRNSM 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GARLLGNPSAGIQS-----TYARLAL------SGGLRHDMGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             696 SSSDQVQAETSSQTTLSPDTTTTSHAPRESSSPPSTSDMLTTTASTEGTSGDTGHTTAVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 -ESKGAGE-------SKGVGELR-------ESNSG--AENTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      66; Mismatches 267;
                                                                                                        5.0%; Score 143.5; 19.5%; Pred. No. 1.4
                                       421 NVDTPGSEDTMESRRSSMASTSSTFFDTSSIG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: mRNA
A; Residues: 1-1630 ewUA>
A; Cross-references: GB: U06746
C; Reywords: glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 145; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DPDAAA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: A53577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          165
```

g

g ò Q õ g δ g ò g ò g ò g ò g ö

à

g

ò

δ

Gaps

```
A;Map position: 2
A;Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Cross-references: EMBL: Z49130; PIDN: CAA88964.1; GSPDB: GN00020; CESP: T06D8.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Apportetical protein T06D8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Dacession: T24583
R;Palmer, S.
submitted to the EMBL Data Library, April 1995
A;Reference number: 219909
A;Reference number: 219909
A;Reference number: translated from GB/EMBL/DDBJ
A;Reference number: bNA
A;Reference bNA
A;Reference iDNA
A;Reference iDNA
A;Residues: 1-1829 <WIL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       559 AISASEEALEDSM-DEGDKKS-----PISQVHEIGIKRNMTVHFKVLREEGPAHMKNFI 611
                                                                                                                                                                                                                                                                                                                                                                                                                         103 IGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKF----VFTGGR 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --EKEPIFELIAKNGNETARREFYMEVSASGSTARGTGNSKK-LAKRNAAQALFEL--- 778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    309 VL------KDDVV---ANIE------EQAKAAGEEAKQQAIENNAQAQKKYDE 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  831 ILILRQNKKPAKKRDQIVIVKSNVESKEEEANKEVAVAAEENSNNSANSGDSSNSSSGDS 890
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347 QQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSAR 406
                                                                                                                                                                                                                                                                                                                        -----PVRNSMADSGDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQ 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    612 TACIVGSIVTEGEGNGKKVSKKRAAEKMLV----ELQKLPPLTPTKQTPLKRIKVKTPGK 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       668 SGAAAREGSVVSGTDGTMQTGKPERRKRLNPPKDKLIDMDDADNPITKLIQLQQTRK--- 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12 VNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNRASDVPGLPV 71 | : | : | : | : | : | : |
                                                                                                                                                                                                                        2 PIG-NLGHNPNVNNSIPPAPPLPSQTDCAGGRGQLINSTGPL-----GSRALFT---- 49
                                                                                                                                                                                                                                                         194 AGESKGVGELRESNSG----AENTTETQTSTSTSSLRSDPKLWLALGTVATGLIGLAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GIVQALALIPEPDSPTITDPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSG
     RNA-binding repeat homology <DSR1>
RNA-binding repeat homology <DSR2>
RNA-binding repeat homology <DSR3>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 115;
                                                                                                                                                                          Indels 107;
                                                                                                                          Length 1026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1829;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.9%; Score 140.5; DB 2;
22.3%; Pred. No. 2.4;
tive 64; Mismatches 191;
                                                                                                                          Query Match
4.9%; Score 140.5; DB 1;
Best Local Similarity 20.0%; Pred. No. 1.1;
Matches 97; Conservative 74; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Experimental source: clone T06D8
     F;308-379/Domain: double-stranded F;575-646/Domain: double-stranded F;708-782/Domain: double-stranded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 106; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Genetics:
A; Gene: CESP: T06D8.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GVHMK 953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      407 TVENK 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local S:
Matches 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   949
                                                                                                                                                                                                                                                                                                                        20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                       Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ۵y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
A; Note: the authors translated the codon GTT for residue 1103 as Tyr
A; Note: part of this sequence, including the amino end of the mature protein, were deter
C; Genetics:
A; Genome: plasmid
C; Superfamily: lactocepin; subtilisin homology
C; Keywords: hydrolase; serine proteinase; transmembrane protein
F; 1-33/Domain: signal sequence #status predicted <SIG>
F; 34-187/Domain: propeptide #status predicted <PRO>
F; 188-1962/Product: serine proteinase, cell-envelope-associated #status predicted <AMT>
F; 208-634/Domain: subtilisin homology #status atypical <SBT>
F; 1938-1955/Domain: transmembrane #status predicted <TMM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: FlyBase:FBgn0003520
C;Superfamily: maternal effect protein; double-stranded RNA-binding repeat homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             egg.
                                                                                                                                                                                                                                                                                                                                                                         20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ....train effect protein (staufen) - fruit fly (Drosophila melanogaster)
C; Species: Drosophila melanogaster
C; Species: Drosophila melanogaster
C; Date: 10.5ep-1999 #sequence_revision 10.5ep-1999 #text_change 10.5ep-1999
C; Accession: A40315
R; Et. Johnston, D.; Beuchle, D.; Nuesslein-Volhard, C.
Cell 66, 51-63, 1991
A;Title: staufen, a gene required to localize maternal RNAs in the Drosophila A; Reference number: A40315; MUID:91300552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1671 VKDALGTDLGNQTDP--STGKTFTAALDDLVAQAQAGTQTDDQLQATLAK-----I 1719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1452 INSGKP-GHMAIDQPVKLLEGKNVLTVAVTDSEDNTTTKNITVYYEPKKTLAAPTVTPS- 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----TTEPAQTVTLTANAAA----TGETVQYSADGGKTYQDVP-----AAGVTIT 1551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KGVGELRESNSGAENTTETQTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TPE---PDSPTTTDPDAAASATETATRDQLTKEA---FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     311 KDDVVANIEEQAKAAGE----EAKQQAI------ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      352 QEELKVSS------GAGYGLSGALILGGGIGVAVTA------ALHRKNOP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    390 VEQTITITITITITSAR-----TVENKPANNIPAQGNVD---TPGSEDIMESRRSSM- 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                         INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIDN: AAA73062.1; PID: 9158506
                                                                                                                                                                                                                                                                                                                                                                         Indels 113;
                                                                                                                                                                                                                                                                                                                        Length 1962;
                                                                                                                                                                                                                                                                                                                      5.0%; Score 141.5; DB 2; 22.8%; Pred. No. 2.3; tive 61; Mismatches 211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 1-1026 <SrJ>
A; Cross-references: GB:M69111; NID:g158505; C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1900 TNPATTTSTTTDDRNG 1918
                                                                                                                                                                                                                                                                                                                        Query Match 5.0
Best Local Similarity 22.8
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----ASTSSTFFDTSSIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Status: preliminary
Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Gene: FlyBase: stau
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     439
                                                                                                                                                                                                                                                                                                                                                                                                                         36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

308

Search completed: September 27, 2001, 14:21:47 Job time: 83 sec

```
NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-K12
                                                                                                                                                                                                                                                                            (without alignments)
1286.019 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          drosophila
haemophilus
rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           erwinia her
pantoea ana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
mesocricetu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pantoea ana
haemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homo sapien
schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           saccharomyc
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       drosophila
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             streptococc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   haemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         equine herp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sus scrofa
                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                  ; Search time 14.89 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                       1 MPIGNLGHNPNVNNSIPPAP......SNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P47035
P12021
P32334 :
P16376 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P98091
P25159
Q09625
Q09625
P16239
Q47879
Q47879
Q47879
Q62346
P16539
P16539
P16539
P16539
P16539
P18669
P1809
P18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P76072
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                              93435 seqs, 34255486 residues
                                                                                  September 27, 2001, 14:22:50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STFR_ECOLI
MCS_BOYIN
STAU_DORONE
US45_LACLC
US45_LACLC
USAB_CABEL
ICEN_PANAN
ICEN_PANAN
ICEN_PANAN
DAB_DROME
SP4_HUMAN
PAP_LACLC
PAP_LUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MSP1_PLAYO
HKR1_YEAST
P1P_LACLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P2P_LACLA
PCX_DROME
IGA4_HAEIN
APC_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APMU_PIG
MSB2_YEAST
7UP2_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BIR1_SCHPO
ICEA_PANAN
IGA0_HAEIN
                                                                                                                                                                                                                                                                                                                                                                Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                _HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEXT_STRDO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JH6_YEAST
                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GA2_HAEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DBA_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /GLX_HSVEB
                                                                                                                                                                                                                  Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                             US-09-189-415A-4
                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                         SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146.5
140.5
136.5
133.5
133.5
130.5
130.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               129
128.5
127.5
127
127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123.5
123
122.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
                                                                                                                                                           Perfect score:
                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                          Jatabase :
                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                          Sequence:
                                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
No.
```

	Q15648 h peroxisom P35820 drosophila P13692 enterococcu				
P60_LISMO P100_HCMVA	PPRB_HUMAN PSC_DROME p54_ENTEC	MUC1_XENLA AKAC_HUMAN	P60_LISSE SLAP_CAUCR	CHI3_CANAL VGP_MABVP	NIT2_NEUCR
			- г		7
484	1581 1603 516	662 1781	523 1025	567 681	1036
4.2	444	4 4	4 4 1.1	4 4 L. L.	4.1
119	119 119 118 5	118	117.5	117	117
34	36 37	9.64	4142	4 4	45

ALIGNMENTS

```
THE STEP.

10. STEP. ECOLI

11. STEP. ECOLI

12. STEP. ECOLI

13. PIGNOZE 1977560.

14. Last sequence update)

15. COTT-2000 (Rel. 40, Last sequence update)

16. COTT-2000 (Rel. 40, Last sequence update)

17. COTT-2000 (Rel. 40, Last sequence update)

18. COTT-2000 (Rel. 40, Last sequence update)

18. STEP. ROLL FILES PROFEIN HONOLOG FROM LAMBDID PROPHAGE RAC.

18. STEP. OR 18. 13. 13.

18. SECRETAL PROFEIN HONOLOG FROM LAMBDID PROPHAGE RAC.

18. Escheriche acoli.

18. Escheriche acoli.

18. SEQUENCE FROM N.A.

19. SEQUENCE FROM N.A.

19. SEQUENCE FROM N.A.

19. SEQUENCE FROM N.A.

10. SECRETAL SECONOR OF THE SUBJECT OF THE SECONOR OF TH
```

us-09-189-415a-4.rsp

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                        DISULFID
                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                            Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
STAU_DROME
                                                                                                                       REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                DOMAIN
DOMAIN
                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                       17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Submarilary gland;
MEDINE=90370871; PubMed=2204065;
MEDINE=90370871; PubMed=2204065;
MEDINE=90370871; PubMed=2204065;
MEDINE=90370871; PubMed=2204065;
Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein containing two distinct domains.";
Proc. Natl. Acad. Sci. U.S.A. 87:6798-6802(1990).
-!- SUBCELLULAR LOCATION: SCENETED.
-!- SIMILARITY: TO PORCINE APOMGUIN.
-!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
                                                                                                                                                                                                                                                                                                                                  356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             416
                                                                                                                                                                                                   124 --RNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRILE 181
                                                                                                                                                                                                                                                                                                    LLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSTSSLRSDPKLWLALGTVATG 241
                                                                                                                                                                                                                                                                                                                                                                                                    242 LIGLAATGIVQALA----LTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVN 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           300 --TAAASSASAASTSAGQASASATAA----GKSAESAASSASTATIKAGEATEQASAAAR 353
                                                                                                                                                     59 SVILLVEGFPPSHAGTITVYEDSQP-GTLNDFLGAM----TEDDARPEALRRFELMVEE 112
                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                                                                                                                                                                                                                                                 113 VARN---ASAVAQNTAAAKKSASDAS-----TSAREAATHA--ADAADSARA----
                                                                                                                                                                                                                                                                                                                                                                                                                                               SAGAAKTSETNASASLQSAATSASTATTKASEAATSARDAAA----SKEAAKSSET---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     298 IDELGNAIPSGVLKDDVVANIEEQAKAAGEEAK-QQAIENNAQAQKKYDEQQAKRQEELK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        252 -----NASSS------ASSAASSATAAGNSAKAAKTSETNAKSSETAAGQSAAAAGSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                357 VSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSARTVENKPANNT
                                                          91;
                                                                                                        SEITLNDGFEVLH------DHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQ----
       Length 1120;
       5.1%; Score 146.5; DB 1; Length 23.1%; Pred. No. 0.39; ive 52; Mismatches 163; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                417 PAQ----GNVDTPGSEDTMESRRSSMASTSSTFFDTSS 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | : || :| || ::| || 354 SASAAKTSETNAKASETSAESSKTAAASSASSASSASSAS 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
01-CT-1996 (Rel. 34, Last annotation update)
SUBMAXILLARY MUCIN-LIKE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     563 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001007; -.
Pfam; PF00007; Cys_knot; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M36192; AAA30657.1;
HSSP; P00214; 1FTC.
Query Match
Best Local Similarity 23.1%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000359; -. InterPro; IPR001007; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MUCS_BOVIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
MUCS_BOVIN
                                                                                                             79
                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                                                         155
                                                                                                                                                                                                                                                                                                                                                                                                                                                       200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DPT TEN SECOND CONTRACTOR OF THE PROPERTY OF T
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                           ò
```

```
13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Drosophila melanogaster (Fruit fly).
Eukaryota, Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SKRSSPGS-KTGNTGAISGTTVAPGSSNTGATTS-----LGSGETTQGGIKIVTMGVTTG 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     254 LALIPEPDSPTTIDPDAAASATETATRDQLIKEAFQNPDNQKVNIDELGNAIPSGVLKDD 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                149 TIIAPGSSNTKATTPTEVRTTTEVRTATETT-----TSRHSSDATGSGIQTGI---- 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            314 VVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGG 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----TGTGSGTTSSPGGFNAEATTFKEH-----VRTTETRILSGTTRGRSGTTVIPE 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     374 GIGVAVTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPANNTPAQGNVDTPGSEDTMES 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        245 SSNTGTSTGVGR-----OTSTAVVSGRVTGV-SESSSPGTSKEASETTTGPGISTTGST 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      135 DQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRILELLEPKGTGESKGA 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ....---QEDGKHIAVGQRNGVETSVVLS 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 KVLQENSPRHAIS---GSSHTEATTLIVSNSTSGTGLRPEDNTAVAGGOATGRVT---- 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  195 GESKGVGELRESNSGA-ENTTETQTSTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQA
                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC. . ) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
AB326CD78E5FFFCF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 142; DB 1; Length 563;
Pred. No. 0.29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.0%; Score 1**,
20.3%; Pred. No. 0.29;
tive 4%; Mismatches 169; Indels
                                                                                                                                                                                                  CTCK.
TO POCCINE APOMUCIN.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=91300552; PubMed=1712672;
St Johnston D., Beuchle D., Nuesslein-Volhard C.;
                                                                                    3 X 11 AA REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
MATENAL EFFECT PROTEIN STAUFEN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1026 AA.
                                                                                                                                                                                 VWFC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88 EVLHDHGPLDTLNRQIGSSVFRVET-
                                                                                                                                                                                                                                                                                                                                                                                                                                58913 MW;
PROSITE; PS01208; VWFC; 1.
PROSITE; PS01185; CTCK_1; 1.
PROSITE; PS01225; CTCK_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                      Glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   434 RRSSMASTSS 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               298 SKSNRITTSS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                              563 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=7227;
                                                                                                                                                                                                       471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STAU_DROME P25159;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              75;
                                                                                                                                                                                                                            SIMILAR
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                          DISULFID CARBOHYD.
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
```

m

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOUR PRESENTATION OF THE PARTY 
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                      ò
                                                                          qq
                                                                                                                          δ
                                                                                                                                                         q
                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                       Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              мау
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ΑT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 between the Swiss Institute of Bioinformatics and the EMBL outstati
the Buropean Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no
modified and this statement is not removed. Usage by and for commer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY: CONTAINS 5 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
Staufen, a gene required to localize maternal RNAs in the Drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----PVRNSMADSGDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQ 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DETERMINANTS TO THE POSTERIOR POLE OF THE DROSOPHILA EGG AND FOR BICOLD RNA TO LOCALIZE CORRECTLY TO THE ANTERTOR POLE.

FOR BICOLD RNA TO LOCALIZE CORRECTLY TO THE ANTERTOR POLE.

THE POSTERIOR POLE.

TISSUE SPECIFICITY: POLAR GRANULES AT THE POSTERIOR POLE OF THE COCYTE, AND BY THE TIME THE EGG IS LAID, AT THE ANTERIOR POLE.

DOMAIN: CONTAINS A PROLINE-RICH DOMAIN. THE INSERTION OF THIS DOMAIN IN THE DRBM 2 DOMAIN IS REQUIRED FOR STAU-OSKAR MRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIG-NLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPL-----GSRALFT---- 49
                                                                                                                          Bycroft M., Grunert S., Murzin A.G., Proctor M., St Johnston D.; "NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of
                                                                                                                                                                                                                                                                                 .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 107;
                                                                                                                                                                                                                                                                                                                                                                MEDLINE-20183617; PubMed-10716936; Micklem D.R., Adams J., Grunert S., St Johnston D.; "Distinct roles of two conserved Staufen domains in oskar mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1026;
                                                                                                                                                                                                                                                                        Bycroft M., Grunert S., Murzin A.G., Proctor M., St Johnston EMBO J. 14:4385-4385(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FUNCTION: REQUIRED BOTH FOR THE LOCALIZATION OF MATERNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inding; Repeat; 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AE5B97624BBF7D0B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.9%; Score 140.5; DB 1;
20.0%; Pred. No. 0.77;
iive 74; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 (ATYPICAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLN/HIS/PRO-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLY-GLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-GLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DRBM 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DRBM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DRBM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DRBM
                                                                                                     MEDLINE-95354674; PubMed-7628456;
                                                                                                                                                                                                                                                                                                                                              CHARACTERIZATION OF DRBM DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                             localization and translation."; EMBO J. 19:1366-1377(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              n; RNA-bi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M69111; AAA73062.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00035; dsrm; 5.
PROSITE; PS50137; DS_RBD; 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1026 AA; 110347
                                                                                   579-646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A40315, A40315.
PDB; 1STU; 31-JUL-95.
FlyBase; FBgn0003520; stau.
InterPro; IPR001159; -.
                                                                                                                                                                                                             EMBO J. 14:3563-3571(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match
Best Local Similarity 20.v..
97; Conservative
                                                                                 BY NMR OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protei
                                    ceil 66:51-63(1991)
                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCALIZATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Developmental
                                                                                 STRUCTURE
                                                                                                                                                                                           ribosomal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                        ERRATUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
```

.. <u>::</u>

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                               194 AGESKGVGELRESNSG-----AENTTETQTSTSSLRSDPKLWLALGTVATGLIGLAAT 248
                                                                                                                                                                                                                                                                                                   --EKEPIFELIAKNGNETARRREFVMEVSASGSTARGTGNSKK-LAKRNAAQALFEL--- 778
                                                                                                                                                                                                                                                                                                                                                                                                   779 --LEAVQVTP-----TNETQSSEECCTSATMSAVTAPAVEATAEGKVPMVATPVGPMPG 830
                                                                                                                                                                                                                                                                                                                                                                                                                                                        VL------KDDVV---ANIE-----EQAKAAGEEAKQQAIENNAQAQKKYDE 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSAR 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | : | ::: | | | : | 891 QATEAASESALNTSTGSNTSGVSSNSGNTDGNNHAESK--NNTESSSNSTSNTQSA 948
                                               103 IGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKF----VFTGGR 158
                                                                                            TACIVGSIVTEGEGNGKKVSKKRAAEKMLV----ELOKLPPLTPTKOTPLKRIKVKTPGK 667
                                                                                                                                               GIVQALALIPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSG 308
559 AISASEEALEDSM-DEGDKKS-----PISQVHEIGIKRNMTVHFKVLREEGPAHMKNFI 611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    χ.
Χ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rutten G., Oteman M., Siezen R.J., de Vos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Cloning of usp45, a gene encoding a secreted protein from Lactococcus lactis subsp. lactis MG1363.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-SER.
51493C42224F3C03 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Rel. 19, Created)
(Rel. 33, Last sequence update)
(Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECRETED 45 KDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 461 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A., AND SEQUENCE OF 28-37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECRETED 45 KDA PROTEIN PRECURSOR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -! - SIMILARITY: TO E.FAECIUM P54.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-MG1363;
MEDLINE-91071599; PubMed-2123812;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47387 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M60178; AAA25230.1; -. EMBL; A17083; CAA01320.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27
461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95:155-160(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1
28. 4
299 3
461 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; JN0097; JN0097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    van Asseldonk M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1359;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GVHMK 953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407 TVENK 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT . US45_LACLC ID US45_LACLC P22865;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1991
01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lactococcus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                 899
                                                                                                                                                                                                                                                                                                   725
                                                                                                                                                                                                                                                                                                                                                                                                                                                           309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sene
```

84306 MW; 76DC5B03E6357A6A CRC64;

```
796 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ICEN_ERWHE P16239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=M1
 SEQUENCE
                                   Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Trwinia
                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
ICEN_ERWHE
                                                                                  89
                                                                                                         137
                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                             ò
  g
                                                                                                         g
                                                                                                                                 ò
                                                                                                                                                        ద
                                                                                                                                                                                δ
                                                                                                                                                                                                       d
                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseéisb-sib.ch).
                          14;
                                                                                                                                                                                                                                                                                          445 FFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDN 504
                                                                                                                                                                                                                                                                                                                                                           | : | : | : | : | : | 338 GSGINSSPIGNPYAGGGCTDYVWQYFAAGGIXIRNIMPGNGGQWASNGPAQGYLHVVGAA 397
                                                                                                                                              ----TKEAFQNPDNQKVNIDE---LGNAIPS------GVLKDDVVANIEEQAKAAG 326
                                                                                                                                                              152 QQEKEQKELSQKSETVKKNYNQFVSLSQSLDSQAGELTSQQAELK---VATLNYQATIAT 208
                                                                                                                                                                                             EEAKQQAIENNAQAQKKYDEQQAKRQE--ELKVSSGAGYGLSGALILGGGIGVAVTAALH 384
                                                                                                                                                                                                                   209 AQDKKQALLDEKAAAEKAAQEAAKKQAAYEAQQKEAAQAQAASTAATAKAVEAATSSASA 268
                                                                                                                                                                                                                                              444
                                                                                                                                                                                                                                                             269 SSSQAPQVSTSTDNTTSNASASNSSNSSSSSSSSSSSSSSSSSSSSNTNS 324
                                                                                                                                                                                                                                                                                                                    ....-TTTG 337
                                                                                                                                                                                                                                                                                                                                           505 GARL----LGNPSAG------IQSTYARLALSG-------GLRHDMG-- 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
NCBL_TaxID=6239;
                                                 181 ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWLALGTVAT 240
                                                                                                241 GLIGL-AATGIVQALALTPEPDSPTTTDPDA-------AASATETATRDQL-- 283
                                                                                                              Gaps
                                                                RKNQPVEQTTTTTTTTTTTSARTVENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSST
   4.8%; Score 136.5; DB 1; Length 461;
19.1%; Pred. No. 0.47;
.ive 69; Mismatches 163; Indels 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
HYPOTHETICAL 84.3 KDA PROTEIN ZK945.10 IN CHROMOSOME II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             796 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane.

POTENTIAL.

SER/THR-RICH.

SER/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; 248582; CAA88469.1; -. EMBL; 248544; CAA88446.1; JOINED. EMBL; 248544; CAA8844.1; -. EMBL; 248582; CAA88444.1; JOINED. WORMPEP; XR945.10; CE01732. Hypothetical protein; Transmembrai
                                                                                                                                                                                                                                                                                                                                                                                             -GLTGGSNSA--VNTSNNP 550
                                                                                                                                                                                                                                                                                                                                                                                                                   398 PGVIASSFSADFVGYANSP 416
                               84; Conservative
                                                                                                                                                                                                                                                                                                                         325 GTSTGNTGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     546
752
                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wilkinson-Sproat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hypothetical
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YS8A_CAEEL
        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                               535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YS8A_CAEEL
                                                                             32
                                                                                                                                                    284
                                                                                                                                                                                                                                                  385
                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                          g
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                           g
                                                                                                                                                   à
                                                                                                                                                                                                   ò
                                                                             염
                                                                                                     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                            14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     291 TSTITVPTSTSTVTTAMSTSTSTPSTSTTTIESTSTFTSTASTSTSTSTTTQCSSSTITS 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -ATGLIGLAATGIVQALALTPEPDSPTTT-----DPDAAASATETATRDQLTKEAFQ 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          290 NPDNQKVNIDELGNAIPSGVLKD--DVVANIEEQAKAAG-EEAKQQAIENNAQAQKKYDE 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      347 QQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTT--S 404
                                                                                                                                                                                                              179 ILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTV 238
                                                                                                                                                                                                                                                                                                                            231 YYDGQVLKGVRAKQFSMRTSGSPTLRRMKRDAGDNTCDYTIESTSTTTTPTTTTVTSTV 290
                                                                                    GLPVNPMRLAASEITILND-----GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHI 119
                                                                                                                                                                                   120 A-VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQR 178
                                                                                                                     GLFLNSTWITLNEVNDDDEISIAVEAKXEVCYDDG----IDRCDGSLWW---LQVGGNEM 189
                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  405 ARTVENKPA-----NNTPAQGNVDTPGSEDIMESRRSSMASTSSTFFDTSSIG 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             406 EPSTSTITIEVIS-------TSSTVTTTEPTTLITSTASTSTIEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Warren G.J., Corotto L.V.; The consensus sequence of ice nucleation proteins from Erwinia herbicola, Pseudomonas fluorescens and Pseudomonas syringae.";
                                            83;
Length 796;
                                               Indels
4.7%; Score 133.5; DB 1;
19.7%; Pred. No. 1.4;
itive 76; Mismatches 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DINET-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1258 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=90152370; PubMed=2515997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ICE NUCLEATION PROTEIN.
                             l Similarity
82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Erwinia herbicola
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=557;
```

```
TGYGSTSTAGADSSLIAGYGSTQTAGYNSILTAGYGSTQTAQENSDLTTGYGSTSTAGYE 701
                                                                                                                                                                                                              PERIODICITY IS SUPERIMPOSED.
                                                                     Biosci. Biotechnol. Biochem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; D14992; BAA03636.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 20.9%
Matches 114; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nterPro; IPR000258; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      266
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                20;
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TLNDGFEVLHDHGPLDT-----LNRQIGSSVFRVETQEDGKHIAVGQRN--- 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               T-TETQTSTSTSSLRSDPKLWLALGTVATGLIG-----LAATGIVQALALTPEPDSPTT- 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOTAGYGSTQTAQKGSD----LTAGYGSTGTAGDDSSLIAGYGSTQ----TAGEDSSLTA 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----TDPDAAASATETATRDQLTKEAF--------QNPDNQK 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296 VNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEEL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DLITGYGSTSTAGA-----NSSLIAGYGSTQTASYNSVLTAGYGSTQTAREGSDL 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KVSSGAGYGLSGAL---ILGGGIGVAVTAALH------RKNQPVEQTTTTTTT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        242 AGYGSTGTAGDDSSLIAGYGSTQTAGEDSSLTAGYGST----QTAQKGSDLTAGYGSTGT 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298 AGADSSLIAG---YGSTQTAGEESTQTAGYGSTQTAQKGSDLTAGYGSTGTAGDDSSLIA 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYGSTQTAGEDSSLTAGYGSTQTAQKGSDLTAGYGSTGTAGADSSLIAGYGSTQTAGEES 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         T----AGYGSTGTAGSDSSIIAGYGSTQTASYHSSLTAGYGSTQTAREQSVLTTGYGSTS 632
                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTTTSARTVENKPANNTPAQGNVDTPGSEDTMESRRSS-----MASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                         TDGAGGRGQLI --- NSTGPLGSRALFTPVRNSMADSGDNRASDVPGLPVNPMRLAASEI - 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TAGADSSLIAGYGSTQTAGYNSILTAGYGSTQTAQEGSDLTAGYGSTSTAGADSSLIAG 691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -GVETSVVLSDQEYARLQSIDPE------GVDKFVFTGGRGGAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----HAMVTVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GYGSTQTAQKGSDLTAGYGSTSTAGYESSLIAGYGSTQTAGYGSTLTAGYGSTQTAQNES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                  Length 1258;
                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                         OCTAPEPTIDE PERIODICITY.

W; 590E8A130077FBD4 CRC64;
                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                              62; Mismatches 226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                  4.6%; Score 132.5; D 20.4%; Pred. No. 2.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1034 AA
                                                                                                                                                                                                                                 Ice nucleation; Repeat; Outer membrane.
DOMAIN 162 1217 OCTAPEPTII
                                                                                                                                 InterPro; IPR000258; -.
Pfam; PF00818; Ice_nucleation; 65.
PRINTS; PR00327; ICENUCLEAIN.
PROSITE; PS00314; ICE_NUCLEAION; 45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Erwinia uredovora)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-94264407; PubMed-7764866;
                                                                                                                                                                                                                                                                             1258 AA; 125084 MW;
                                                                   EMBL; M26382; AAA24823.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ICE NUCLEATION PROTEIN INAU
                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 20.45
Matches 110; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                        PIR; JQ0188; JQ0188.
HSSP; P06620; 1INA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=553;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pantoea ananas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-KUIN-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ICEN_PANAN
Q47879;
                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pantoea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ICEN_PANAN
ID ICEN_PANAN
ID ICEN_PANAN
DT 01-NOV
DT 01-NOV
DT 01-OCT
DT 01-OCT
DT 00-OCT
DT 00-OCT
DT 00-OCT
DT 01-OCT
DT 01-NOV
DT 01-N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              356
                                                                                                                                                                                                                                                                                                                                                                                                                                         56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                               A-G-Y-G-S-T-X-T; FURTHER ON A 16-RESIDUE AND A REGIONAL 48-RESIDUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------HAMVTVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAEN 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----QNPDNQK 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GYGSTQTAQXGSDLTAGYGSTSTAGYESSLIAGYGSTQTAGYGSTLTAGYGSTQTAQNES 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VNIDELGNAIPSGVLKDDVVANIEEQAKA-----AGEEAKQQAIENNAQAQKKYDEQQ 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----TLNDGFEVLHDHGPLDT-----LNRQIGSSVFRVETQEDGKHIAVGQRN--- 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           242 AGYGSTGTAGDDSSLIAGYGSTQTAGEDSSLTAGYGST----QTAQKGSDLTAGYGSTGT 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYGSTQTAGEDSSLTAGYGSTQTAQKGSDLTAGYGSTGTAGADSSLIAGYGSTQTAGEES 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOTAGYGSTOTAOKGSD----LTAGYGSTGTAGDDSSLIAGYGSTQ----TAGEDSSLTA 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AKRQEELKVSSG----AGYGLSGALILGGGIGVAVTAALHR-----KNQPVEQTTTTT 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AQENSDLTTGYGSTSTAGYDSS----LIAGYGSTQTAGYHSILTAGYGSTQTAQERSDLT 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T---TTTTTTSART--VENKPANNTPAQGNVDTPGSEDTMESRRSS-----MASTSSTFFD 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    183 TETAGDSSTLIAGYGSTGTAGSDSTLVAGYGSTQTAGE-ESSQMAGYGSTQTGMKGSDLT 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AGADSSLIAG---YGSTQTAGEESTQTAGYGSTQTAQKGSDLTAGYGSTGTAGDDSSLIA 354
                                                                                                                                                                                                                                                                                                                                                                                                                            -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.-!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
                                             Erwinia
                                                                                                                              58:762-764(1994).
PROTEINS ENABLE BACTERIA TO NUCLEATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GVETSVVL,SDQEYARLQSIDPE-------GKDKFVFTGGRGGAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     T-TETQTSTSTSSLRSDPKLWLALGTVATGLIG-----LAATGIVQALALTPEPDSPTT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DLITGYGSTSTAGA-NSSLIAGYGSTQTASYNSVLTAGYGSTQTAREGSDLTAGYGSTQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TDGAGGRGQLI - - - NSTGPLGSRALFTPVRNSMADSGDNRASDVPGLPVNPMRLAASEI -
                                                                                                                                                                                                                                                                                                          A CONSENSUS OCTAPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1034;
                                        gene of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ice nucleation; Repeat; Outer membrane.
DOMAIN 162 993 OCTAPEPTIDE PERIODICITY.
SEQUENCE 1034 AA; 103378 MW; FA222523D333EADD CRC64;
fichigami Y., Watabe S., Abe K., Obata H., Arai S., "Cloning and sequencing of an ice nucleation active
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------TDPDAASATETATRDQLTKEAF------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.6%; Score 130.5; DB 1;
20.9%; Pred. No. 2.9;
ive 64; Mismatches 229;
                                                                                                                                                                                                           CRYSTALLIZATION IN SUPERCOOLED WATER. SUBCELLULAR LOCATION: OUTER MEMBRANE. DOMAIN: CONTAINS IMPERFECT REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00818; Ice_nucleation; 51.
PRINTS; PR00327; ICENUCLEATN.
PROSITE; PS00314; ICE_NUCLEATION; 34
```

18.98;

```
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                       g
                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                              δ
                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is n no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-93041287; Pubmed=1419850; Maier D., Stumm G., Kuhn K., Preiss A.; Maier D., Stumm G., Kuhn K., Preiss A.; Mairless, a Drosophila gene involved in neural development, encodes a novel, serine rich protein."; Mech. Dev. 38:143-156(1992).

1- FUNCTION: IS A POTENT ANTAGONIST OF NEUROGENIC GENE ACTIVITY IS RESORY ORGAN DEVELOPMENT. THE EXPRESSION OF DISTINCT CELL FATES BY THE TRICHOGEN (SHAFT) / TORMOGEN (SOCKET) SISTER CELL PAIR DEPENDS ON THE LEVEL OF H ACTIVITY. A CERTAIN THRESHOLD LEVEL OF H ACTIVITY IS REQUIRED, BELOW WHICH BOTH SISTER CELLS ADOPT THE TRICHOGEN (SOCKET) SISTER CELLS ADOPT THE TRICHOGEN (SOCKET) SISTER CELLS ADOPT THE TRICHOGEN (SOCKET) SISTER CELLS ADOPT THE TRICHOGEN STREET 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-92387549; PubMed-1516831;
Bang A.G., Posakony J.W.;
"The Drosophila gene Hairless encodes a novel basic protein that "The Drosophila gene Hairless encodes a novel basic protein controls alternative cell fates in adult sensory organ development."; Genes Dev. 6:1752-1769(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
TISSUE SPECIFICITY: OVARY, EMBRYOS, LARVAL AND PUPAL IMAGINAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MISSING (IN REF. 2).
fw; A94BF1A27579E2F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A -> R (IN REF. 2).
AAVA -> RLLP (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HIS/PRO-RICH (PRD MOTIF)
S -> A (IN REF. 2)
QH -> LL (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein; DNA-binding
                                                                                                                                                                                                                                                                                                                           01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
01-FEB-1994 (Rel. 28, Last annotation update)
                                                                                                                                                                                                                                                                1077 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-ALA.
POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY: CONTAINS A "PRD MOTIF"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALA-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rhr-rich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X67239; CAA47664.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | protein; Nuclear pl
| 123 | THR. | 123 | POLN | 142 | POLN | 137 | 946 | POLN | 137 | 946 | POLN | 154 | 1008 | HIS/ | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 | 151 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FlyBase; FBgn0001169; H.
                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P04002; 1WFA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           702
891
964
974
1077
                                                                                                                                                                                                                                                                                                                                                                                                                                HAIRLESS PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ephydroidea; Dro:
NCBI_TaxID=7227;
                                                       448 TSSIGG 453
                                                                                        :| | |
702 SSLIAG 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Developmental
                                                                                                                                                                                                                                                                             HLES_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                           RESULT 8
HLES_DROME
                                                                                                                                                                                                                                                                             ò
                                                                                                                  엄
```

Score 130; DB 1; Length 1077;

4.68;

Query Match

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463 TSEDEEPMTELPRITUAV-NGDLNGDLKASIGKPKSKPKPKAKLSSI----1QKLIDSV 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         403 TNEIKIEKPDTIKGEDDAERLEKEPKKAVSDDSESKEASPGQQVEPQPKDETVDVEMKMN 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            298 IDE------LGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQ 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      247 ATGIVQALALTPEPDSPTTTDPD-----------AAASATE---TATRDQL 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            345 A--LRTAARKRRPHEPLTISEDQQPIFATAIKAENGDDTLKAEAAEAVEIENVAVADTT 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EGKDKFVFTGGR--GGAGHAMVTVASDITEARQRILELLEPKGTGESKGAGESKGVGELR 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----NOKVN 297
                                                                                                                                                                                                                                                                                                                                                                                                                                          EVLHDHGPLDTL-NRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDP 146
                                                                                                                                                                                                                                                                                                                                                   152 AAVAASVVGATASKPTIDVLGGVLDYSSLGGAATGSLPTTAVVAAAAGTAKIGKGSNSGG 211
                                                                                                                                                                         Pred. No. 3.2;
; Mismatches 244; Indels 312; Gaps
                                                                                                    -----LPSQTDGAGGRGQLINSTGPLGS--RALF 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  348 QAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |: :: | |: :: | |: 517 PARLEQMSKTSAVIASTTTSSDRIGGGLSHALT.--HKVSPPSSATAAGRLVEYHTQHVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             574 PRKRILREFEKVSLEDNGCVNNGSGGASSGGAGGKRSRAKGTSTSSPAGKASPMNLAPPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GNVD-TPGSE------DTMESRRSSM-----ASTSSTFFDTSSIGGPCRIRMLM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          634 GKPSPSPGSSSSSTSPATLSTQPTRLNSSYSIHSLLGGSSGSGSSSSSSGKKC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             462 LKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----TTTTTTTSARTVENKPANNTPAQ
                                                                                                                                                                                                                                                                         49 TPVRNSMADSGDNRAS-DVPG------LNDGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----KLWLALGTVATGLIGLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  522 RLALSGGLRHDMGGLTGG----SNSAVNTSNNPPAPGS-----HRF 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            715 RALLTSPKSPDVSGSNGGGGKSPSHTGTKKRSPPYSAGSPVDYGHSF 761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGWVSVTRKTFRPPSAATSATVTPTSAVTTA-----YPKN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        284 TKE-AFQNPD-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-FEE-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
DISABLED PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ŗ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ESNSGAENTTETQTSTSTSSLRSDP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE=93194063; PubMed=7680635;
                                           :99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                               Matches 145; Conservative
                                                                                                                                  7 GHNPNVNNSIPPAPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=7227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DAB_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P98081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DAB_DEROME_

1D DAB_DAB_DAB_DAC

AC 01-FE DT 01-
```

ė,

```
VLKDDVVANIEEQAKAAGEEA----KQQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
SP4_HUMAN
                                  399
                                                                                               459
                                                                                                                                                           519
                                                                                                                                                                                         437
                                                                                                                                                                                                                          555
                          g
                                                               á
                                                                                           q
                                                                                                                          δ
                                                                                                                                                     g
                                                                                                                                                                                         ò
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVET 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||::|::|::|:::|
PVHKISFIAQDMTDSRAFGYI--FGSPDS----GHRFFGIKTDKAAAQVVLAMRDLFQV 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SVVLSDQE--------159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167 VFELKKKEIEMARQQIQGKSLHDHSSQLASLSSLKSSG-----LGGMGLGHSDLASGGI 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAGHAMVTVASDI--TEARQRI---LELLEPKGTGESKGAGES-----KGVGELRES 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     221 SSGHALTLIGSSLSTINGTSRLGVSLDVAKASGSAAKEVSPESVADLVDLEQELTSLQRG 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NSGAENTTETQTSTSTSSLRSDPKLWLalGTVATGLIGLAATGIVQALALTPEPDS---- 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | | | | : :||::
ISQMERITPNEPTTSSTGGAGHPS--LAKSASEDDPFGDSFIYVPSYSILPPPPDSGRNR 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --PTTTDPDAAASATE-----TATRDQLTKEAFQNPDNQKVN----IDELGNAI-PSG 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.5%; Score 129.5; DB 1; Length 2411;
18.8%; Pred. No. 9.7;
Live 79; Mismatches 228; Indels 199; Gaps
                                    ಥ
                                                                                                                              K
                                                                                                  DEVELOPMENT.

1 SUBCELLULAR LOCATION: CYTOPLASMIC.

1 ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND 1 SHORT FORM; ARE PRODUCTED BY ALTERNATIVE SPLICING.

1 TISSUE SPECIFICITY: CNS AXONS AND BODY WALL MUSCLES.

1 DEVELOPMENTAL STAGE: EMBRYONIC AXONOGENESIS.

1 PTM: PROBABLY PHOSPHORYLATED BY THE ABL TYROSINE KINASE.

1 SIMILARITY: CONTAINS 1 PID DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
PHOSPHORYLATION (BY ABL) (PROBABLE).
Gertler F.B., Hill K.K., Clark M.J., Hoffmann F.M.;
"Dosage-sensitive modifiers of Drosophila abl tyrosine kinase function: prosperor, a regulator of axonal outgrowth, and disabled, novel tyrosine kinase substrate.";
Genes Dev. 7:441-453(1993).
-1- FUNCTION: TOGETHER WITH ABL INVOLVED IN EMBRYONIC NEURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESIDUE.
RESIDUE.
RESIDUE.
(PROBABLE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MISSING (IN SHORT ISOFORM). W; 579AB9C0243D5FD6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT-RICH REGION,
ALTERNATE ARG AND ACIDIC
ALTERNATE ARG AND ACIDIC
ALTERNATE ARG AND ACIDIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PHOSPHORYLATION (BY ABL) PHOSPHORYLATION (BY ABL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Phosphorylation
                                                                                                                                                                                                                                                                                                                                                                   EMBL; L08845; AAB08527.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               264046
                                                                                                                                                                                                                                                                                                                                                                                                 FlyBase; FBgn0000414; Dab
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 18.8%
Matches 117; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000050; -.
Pfam; PF00640; PID; 1.
PROSITE; PS01179; PID; 1.
Alternative splicing; Pho
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      482
1662
1667
1701
1704
1713
1739
1826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1713 171
1739 173
1826 183
1961 196
462 67
                                                                                                                                                                                                                                                                                                                                                                                     A46299; A46299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111
482
1662
1667
1701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOD_RES
MOD_RES
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   263
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                 SLDAFTDLDPLGTGRTRPYVDKKYFFQELKNPPKKLLKELSSGSQAGLGLGLSLGQLDGL 518
                                                                                                                                                                                                                                                                                               -----SGGLRHDMGG 535
                                                                                                                                                                                                                                                                                                                     VVSSSGLGSVLAMAPLASSESTATPTQQLTEVAAGSGPLADLDIGLSTALGNEEQTSTIL 458
                                                                                                                         AVTAALHRKNQPVEQTTTTTTTTTTTSARTV-ENKPANNTPAQGNVDTPGSEDTMESRRS 436
                                                                                                                                                                 -----RKK 554
                                                                                                                                                                                                          SMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQH 496
                                                                                                                                                                                                                                                   SLTT------EMHILYYDKRVVHFWRNFFSVQLEIALSKQLSKVCTCTAG 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ozersky P., Holmes A.;
Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: BINDS TO GT AND GC BOXES PROMOTERS ELEMENTS. PROBABLE
TRANSCRIPTIONAL ACTIVATOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hagen G., Mueller S., Beato M., Suske G.;
"Cloning by recognition site screening of two novel GT box binding
proteins: a family of Spl related genes.";
Nucleic Acids Res. 20:5519-5525(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: NUCLEAR.
TISSUE SPECIFICITY: ABUNDANT IN BRAIN.
SIMILARITY: BELONGS TO THE SPI FAMILY OF C2H2-TYPE ZINC-FINGER
                                         --AQKKYDEQQAKRQEE---LKVSSGAGYGLSGALILGGGIGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                 F-----PEDSTTISTTTTTATNITAVLTNRYSNTIIAQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SP4_HUMAN STANDARD; PRT; 784 AA. 002446: 060402; 01-FBB-1995 (Rel. 31, Created) 01-FBB-1995 (Rel. 31, Last sequence update) 01-OCT-2000 (Rel. 40, Last annotation update) TRANSCRIPTION FACTOR SP4 (SPR-1).
                                                                                                                                                                                                                                                                                               497 PPRDITDNGARLLGNPSAGIQSTYARLAL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-93087156; PubMed-1454515;
Hagen G., Mueller S., Beato M., S
                                                                                                                                                                                                                                                                                                                                                                                536 LTGGSNSAVNTSNNPPAPGSHRF 558
                                                                                                                                                                                                                                                                                                                                                                                                        : | | | | : | VT------NPPNPTGHYY 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X68561; CAA48563.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AC004595; AAD12226.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MIM; 600540; -.
InterPro; IPR000822; -.
Pfam; PF00096; zf-C2H2; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; S26638; S26638.
HSSP; P08047; 1SP1.
TRANSFAC; T02339; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rissue=Uterus;
```

```
Local Simhes 121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               111
                                                                                                                                                                                                                                                                <u>|</u>.
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                        21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---KAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAV 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    459 PTLTPSGQISWQTVQVQNIQSLSNLQVQNAGLSQQLTITPVSS---SGGTTLAQIAPVAV 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        380 TAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPANNTPAQGNVDTPGS-----EDTME 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGA-----PITLNTAQLASVPNLQTVSVANLGAAGVQVQGVPVTITSVAGQQQGQDGVK 569
                                                                                                                                                                                                                                                                                                                                                                                                                        246 AATGIVQALALIPEPDSPTTTDPDAAASATETATR--DQLTKEAFQNPDNQKVNIDELG- 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----- 322
                                                                                                                                                                                                                                                                                                                                                                                   193 GAGESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWLALGTVAT-----GLIGL 245
                                                                                                                                                                                                                                                                                                                                                                                               342 ADIG---QYASTSASSERTIEESQTPAATESEAQSSSQLHANGMONQQDQSNSLQOVQI 398
                                                                                                                                                                                                                                                                                                       108 FRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFT-GGRGGAGHAMV 166
                                                                                                                                                                                                                                                                                                                                                               222 ONLANOTVPVQIRPGVSIPLQLQTLPGTQAQVVTTLPINIGGVTLALPVINNVAAGGGTG 281
                                                                                                                                                                                                                                              58 IGTPGENQATGQQQIIIDPSQGLVQLQNQPQQLELVTTQLAGNAWQLVASTPPASKENNV 117
                                                                                                                                                                                                                                                                 48 FTPVRNSMADSGDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV 107
                                                                                                                                                                                                                                                                                    SQPASSSSSSSSNNGS----ASPTKTKSGNSSTPGQFQVIQVQNPSGSVQYQVIPQL 171
                                                                                                                                                                                                                                                                                                                           221
                                                                                                                                                                                                                                                                                                                                              -----GESK 192
                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                          3 IGNLGHNPNVNNS---IPPAPPL-----PSQ-----TDGAGGRGQLINSTGPLG-SRAL 47
                                                                                                                                                                                                                                                                                                                   172 QTVECQQI-----QINPTSSSSLQDLQCQIKLIS---AGNNQAILTAANRTASGNILA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TI-1996 (Rel. 34, Created)
TI-1996 (Rel. 34, Last sequence update)
TI-2000 (Rel. 40, Last annotation update)
CHL. PACTOR CI (HCF) (VP16 ACCESSORY PROTEIN) (HFCL) (VCAF)
         PROSITE; PSO0028; ZINC_FINGER_C2H2_1; 3.
PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
Transcription regulation; Activator; Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat.
                                                                                                                                                                                       Query Match
4.5%; Score 129; DB 1; Length 784;
Best Local Similarity 20.0%; Pred. No. 2.5;
Matches 112; Conservative 71; Mismatches 234; Indels 144;
                                                                                                                                                                                      Length 784;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------VLKDDVVANIEEQA-
                                                                                                                                                  -> A (IN REF. 2).
3C4EAE28CB2B81FB CRC64;
                                                                                                                             QP (IN REF. 2).
QP (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 2090 AA
                                                                                                                                                                                  Score 129; DB
                                                                                                                                                                                                                                                                                                                                                -----TVASDITEARQRILELLEPKGT-----
                                                                               POLY-SER.
ZINC FINGERS.
                                                                                                   C2H2-TYPE.
C2H2-TYPE.
                                                                                                                      C2H2-TYPE
                                                            POLY - ALA.
                                                    POLY-GLU.
                                                                       POLY-SER
                                                                                                                                o
^-
                                                                                                                                K -> 0
HA ->
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433 SRRSSMASTSSTFFDTSSIGG 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     570 VOOATIAPV-----TVAVGG 584
                                                                                                                                                            Ψ.
      PR00048; ZINCFINGER
                                                                                                                                                            82025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                             701
729
197
386
                                                                       130
188
729
                                                                                                                                                    386
784 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-OCT-1996 (
01-OCT-1996 (
01-OCT-2000 (
HOST CELL FAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HFC1_MESAU
P51611;
                                                                                                                                          CONFLICT
                                                                                                                                                              SEQUENCE
                                                                                 DOMAIN
DOMAIN
ZN_FING
ZN_FING
ZN_FING
                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CFF).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
HFC1_MESAU
                                                      DOMAIN
                                                                         DOMAIN
                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                  167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                  ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNRASDV-----PGLPVNPMRLAASEI---TLNDGFEVLHDHGPLDTLNRQIGSSVFRV 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ETQEDGKHIAVGQRNGVETSVVLSDQEY-----ARLQSIDP-----EGKDKFVF.154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         729 DGKPTTIITTTQASGAGSKPTILGISSVSPSTTKPGTTTIIKTIPMSAIITQAGAT---- 784
                                                                                                                                                                                                                                                                                                                                                                                                                       SUBDINIT: COMPOSED PREDOMINANTLY OF SIX POLYPEPTIDES RANGING FROM 110 TO 150 KDA AND A MINOR 300 KDA POLYPEPTIDE. THE MAJORITY OF NAND C-TERMINAL CLEAVAGE PRODUCTS REMAIN TIGHTLY, ALBEIT
                                                                                                                                                                                  GOLO H., Nishitahi H., Umene K.I., Nakabepu Y., Nishimoto T.;
Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: UPON LYTIC INFECTION OF PERMISSIVE CELLS, THE HSV
-!- FUNCTION: UPON LYTIC INFECTION OF PERMISSIVE CELLS, THE HSV
--- ARANGACTIVATOR PROTEIN VPI6 ASSOCIATES WITH HCF. BINDING TO HCF
ACTIVATES VPI6 FOR ASSOCIATION WITH THE OCTAMER MOTIF-BINDING
PROTEIN OCT-1, TO FORM A MULTIFROTEIN-DNA COMPLEX RESPONSIBLE FOR ACTIVATING TRANSCRIPTION OF THE HSV IMMEDIATE EARLY GENES (BY STANLARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 IGNLGHNPNVNNSIPPAPPLPSQTDGAGGRG---QLINSTGPLGSRALFTPVRNSMADSG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NONCOVALENTLY, ASSOCIATED (BY SIMILARITY).
SUBCELLULAR LOCATION: NUCLEAR.
DOMAIN: THE HCF REPEAT IS A HIGHLY SPECIFIC PROTEOLYTIC CLEAVAGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL (BY SIMILARITY).
PTM: EXISTS AS A FAMILY OF POLYPEPTIDES THAT ARISE THROUGH
CLEAVAGE AT ONE OR MORE CENTRALLY LOCATED SITES. CLEAVAGE AT A DEFINED SITE, PPCE/THET, WITHIN THE HCF REPEAT (BY
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71; Mismatches 211; Indels 224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCF REPEAT 1.
HCF REPEAT 1.
HCF REPEAT 3.
HCF REPEAT 3.
HCF REPEAT 4.
HCF REPEAT 5.
HCF REPEAT 5.
HCF REPEAT 6.
HCF REPEAT 6.
HCF REPEAT 6.
HCF REPEAT 7.
HCF REPEAT 8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 2090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E495E8B1F2385E17 CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 19.3%; Pred. No. 8.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY: CONTAINS 5 KELCH REPEATS.
Mesocricetus auratus (Golden hamster).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KELCH 4.
KELCH 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           214942 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D45419; BAA08258.1; -. HSSP; P02751; 1FNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interpro; IPR00177; -. Pfam; PF00041; fn3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Repeat
                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nuclear protein;
REPEAT 44
                                                                                                               NCBI_TaxID=10036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        266
1010
1010
1072
1107
11157
11295
11323
1323
1423
```

4

```
DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                 SIGNAL
                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
TRX_DROME
                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      257
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                   -----TSTATTAMS----SMPCETHETGT-----TSTATTAMS----SMG 1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
PIII-TYPE PROTEINASE PRECURSOR (EC 3.4.21.96) (LACTOCEPIN) (CELL WALL-
                                                                                                                                                                                                                          ------EAQPVHDLPVSILASPTTEQPTATVTIADSGQGDVQPG------ 1007
                                                                                                                                                                                                                                                                                 --IVILVCSNPPCETHETGTINIATIT--VVANLGGHPQPTQVQFVCDRQEAAASLVTSA 1063
                                                                                                                                           DPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAG 326
                                                                                                                                                                                                                                                                                                                ----GNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRLL 474
----GVTSTPGIKSPITIITTKVMTSGTGAPAKIITAVPKIATGHGQQGVTQVVL 835
                             TGGRGGAGHAMVTV-----ASDIT-EARQRILELLEPKGTGESKGAGESKGVGELRES 206
                                                                                                                                                                                                327 EEAKQQAIENNAQAQKKYD-----EQQAKRQEELKVS---SGAGYGLSGALILGGGIGV 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VOS P., SIMONS G., SIEZEN R.J., de VOS W.M.;

"Primary structure and organization of the gene for a procaryotic, cell envelope-located serine proteinase.";

J. Biol. Chem. 264:1379-13585(1989).

-!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE GROWTH OF THE BACTERIA ON MILK.

-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.

-!- SUBLIBLIARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
                                            NSGAENTTETQTSTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTT
                                                                                                             896 GAGAHSTSASLATPITT-----LGTIAT----LSSQVINPTAITV-----
                                                                                                                                                                                                                                                                                                                                           1064 VGQQNGNVVRVCSNPPCETHETGTTNTATT--ATSNMAG-----QHGC-----
                                                                                                                                                                                                                                                                                                                                                                        475 ILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris)
                                                                                                                                                                    ---SAAQTTLTAAGGLTTPTITMQPVSQPTQVTLI--TAPSGV----
                                                                                                                                                                                                                                                       378 AVTAALHRKNQPVEQTTTTTTTTTTSARTVENKPANNTPAQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1902 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 188-197
                                                                                                                                                                                                                                                                                                                                                                                                                                                          -- TGQQRDTRHTSSNPTVVRITVAPGA 1153
                                                                                                                                                                                                                                                                                                                                                                                                                              535 GLIGGSNSAVNTSNNPP-----APGS 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; J04962; AAA03533.1; ALT_SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-89340435; PubMed-2760036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ASSOCIATED SERINE PROTEINASE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBTILASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-1359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P3P_LACLC
P15292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmid
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1129
                            155
                                                        936
                                                                                                                                         267
                                                                                                                                                                                                                                                                                    1008
                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                   1105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LACLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                     셤
                                                                                                                                     δ
                                                                                                                                                                  qq
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                       ð
                          ò
                                                                                  ò
                                                                                                          g
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
1780 SLKTKVAAAVEAAKTVGKGDGTTGTSDKGGGQGTPAPPAPGDTGKDKGDEGSQPSSGGNIP 1839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EXTRACELLULAR (POTENTIAL).

CYMBRANE ANCHOR (POTENTIAL).

CYTOPLASMIC (POTENTIAL).

CHARGE RELAY SYSTEM (BY SIMILARITY).

CONSERVED IN GRAM-POSITIVE COCCI SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1452 INSGKP-GHMAIDQPVKLLEGKNVLTVAVTDSEDNTTTKNITVYYEPKKTLAAPTVTPS- 1509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----TIEPAQTVTLTANAAA----TGETVQYSADGGKTYQDVP------AAGVTIT 1551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTTALAAATQKAQ-TALDQTNASVDSLTGANRDLQTAINQLAAKLPADKKTSLLNQLQS 1670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1671 VKDALGTDLGNQTDP--STGKTFTAALDDLVAQAQAGTQTDDQLQATLAK-----I 1719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1720 LDEVLAKLAEGIKAATPAEVGNAKDAATGKTWYADIADTLTSGQASADASDKLAHLQALQ 1779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSID 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   146 PEGKDKFVFTGGRGGAGHAMVTVASDI-----TEARQRILELLEPKGTGESKGAGES 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KGVGELRESNSGAENTTETQTSTSTSSLR-SDPKLWLALGTVATGLIGLAATGIVQALAL 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TPE---PDSPTTTDPDAAASATETATRDQLTKEA----FQNPDNQKVNIDELGNAIPSGVL 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --ENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QEELKVSS------ALHRKNQP 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.5%; Score 128.5; DB 1; Length 1902;
22.7%; Pred. No. 8.1;
ive 56; Mismatches 197; Indels 105; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36 INSTGPLGSRALFTPVR-----NSMADSGDNRASDVPGLPVNPMR-LAASEITLND 85
                                                                              Pfam; PF00746; Gram_pos_anchor; 1.
Pfam; PF00082; Peptidase_S8; 3.
Pfam; PF00082; SubTILIS.
PROSTIE; PS00137; SUBTILIASE_ASP; 1.
PROSTIE; PS00137; SUBTILASE_BRS; 1.
PROSTIE; PS00138; SUBTILASE_BRS; 1.
PROSTIE; PS001343; STAM POS_ANCHORING; 1.
PROSTIE; PS00343; GRAM POS_ANCHORING; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87CECBAA9345F9D3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : |||:|| | | | : : | | 1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      390 VEQTITITITITITARIVENKPANNIPAQGNVDIPGSEDIME 432
                                                                                                                                                                                                                                                                                                                                                                                                                                           PIII-TYPE PROTEINASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRX_DROME STANDARD; PRT; 3726 AA. P20659; Q27255; Q27327; 01-FEB-1991 (Rel. 17, Created) 01-OCT-2000 (Rel. 40, Last sequence update) 01-OCT-2000 (Rel. 40, Last annotation update) TRITHORAX PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KDDVVANIEEQAKAAGE----EAKQQAI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1902 AA; 200550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 22.7
Matches 105; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1876
1895
1902
217
                      InterPro; IPR000209; -
MEROPS; S08.019; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    217
281
620
1867
                                                                                                                                                                                                                                                                                                                                                ransmembrane
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                             MEDINE-95009521; PubMed-7924996; Sedkov Y., Tilib S., Mizrokhi L., Mazo A.; Sedkov Y., Tilib S., Mizrokhi L., Mazo A.; The bithorax complex is regulated by trithorax earlier during brosophila embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early trithorax
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -i- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-i- MISCELLANEOUS: THIS PROTEIN HAS BEEN EXPERIMENTALLY SHOWN TO BIND
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                       Mazo A.M., Huang D.-H., Mozer B.A., Dawid I.B.;
"The trithorax gene, a trans-acting regulator of the bithorax complex in Drosophila, encodes a protein with zinc-binding domains.";
Proc. Natl. Acad. Sci. U.S.A. 87:2112-2116(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genes Dev. 8:2478-2490(1994).
--- FUNCTION: FUNCTIONS IN SEGMENT DETERMINATION THROUGH INTERACTION WITH GENES OF BITHORDAX (BX-C) AND ANTENNAPEDIA (ANT-X) COMPLEXES.
IT CAN BEHAVE AS AN ACTIVATOR OF BX-C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-95047388; PubMed-7958911; MEDLINE-95047388; PubMed-7958911; Mazokhi L., Mazo A.; Kuzin B., Tillib S., Sedkov Y., Mizrokhi L., Mazo A.; "The Drosophila trithorax gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene fork head."; Genes Dev. 8:248-2490(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SIMILARITY: BELONGS TO THE TRITHORAX FAMILY OF TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tillib S., Sedkov Y., Mizrokhi L., Mazo A.; "Conservation of structure and expression of the trithorax gene between Drosophila virilis and Drosophila melanogaster."; Mech. Dev. 53:113-122(1995).
                                                                                                                                                                                                                                                       SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1- SIMILARITY: CONTAINS 1 'SET' DOMAIN.
-1- SIMILARITY: CONTAINS 5 PHD ZINC-FINGER DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SUBCELLULAR LOCATION: NUCLEAR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-96100387; PubMed-8555104;
                                                                                                                                             MEDLINE-90192757; PubMed-2107543;
                                                                                                                                                                                                                                                                                                                                                                                       Development 120:1907-1917(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001214; -.
InterPro; IPR001965; -.
Pfam; PF00628; PHD; 2.
Pfam; PF00856; SET; 1.
PROSITE; PS50016; PHD; 3.
PROSITE; PS50280; SET; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FlyBase; FBgn0003862; trx.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; A35085; A35085
                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T00850
                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-OREGON-R;
                                                                                                                                                                                                                                                                                                                                                                             transcripts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRANSFAC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP;
```

```
POLY-SER.
POLY-SER.
POLY-SER.
POLY-SER.
GLN-RICH.
GLN-RICH.
POLY-GLU.
POLY-GLU.
POLY-GLU.
MISSING (IN SHORT ISOFORM).
P -> PWITSPLKFLGLSTHGGLLWLLGVVVRLKQGG (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1048 GVPQAELAGETNATGDTLKRQRIDLP------GP-----RVKHVC---RSAS 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1086 IVLGQPLATFGEDQQPEDAADMQQEIAAPVPSAIMEPSPEKPTHIVTDENDNCASCKTSP 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           962 K------ISNAADQKLAAARAISPSLTKKNSK-QEKEKVKESEQSEK 1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          415 NTP-AQGNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKHRCMIRRCRL 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             474 LILIRLFRIWG------IQISV---VYSTIQHPP------RDTTDNGARLLGNP 512
                                                                                                                                                                                                                                                                                                                                                                          176 RQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWLAL 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          295 KVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEE 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            355 LKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTTSARTVENKPAN 414
                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Streptococcus downei (Streptococcus sobrinus).
Plasmid pYA902.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
                                                                                                                                                                                                                                                                                                                                                                                            236 GIVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEA-FQNPDNQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  923 GDSTSAL-----TSIKPNPLAENNVTFGSTPLLRPAILENPLFL
Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                    Indels 113;
                                                                                                                                                                                                                                                                                                                         Length 3726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1146 VGDESKPSKSSGSAQAEVKKATALG-KEGTASAAGGSSAKVTTRN 1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
DEXTRANASE PRECURSOR (EC 3.2.1.11) (ALPHA-1,6-GLUCAN-6-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----SAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNTSN 548
                                                                            (ATYPICAL)
                                                                                                                                                                                                                                                                -> S (IN REF. 1).
-> S (IN REF. 1).
D2756E50763DICF5
               Developmental protein; Activator;
                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                     148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ą.
                                                                                                                                                                                                                                                                                                                             4.5%; Score 127.5; I
22.0%; Pred. No. 22;
Live 55; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1337
                                       PHD-TYPE 1.
PHD-TYPE 2.
PHD-TYPE 3.
PHD-TYPE 4 (
                                                                                                      SET DOMAIN
                                                                                                                POLY-SER
                                                                                                                                                                                                                                                                                           AA; 400575
                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                          3040
                            Alternative splicing
                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 89; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLUCANOHYDROLASE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEXT_STRDO P39653;
                                                                                                                                                                                                                                                                               CONFLICT
SEQUENCE
                                                                                                                                                                                                                                ARSPLIC
                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                             CONFLICT
                   Nuclear
                                                       ZN_FING
                                                                                ZN_FING
                                                                                          ZN_FING
                                                                                                                                             DOMAIN
                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEXT_STRDO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
BIR1_SCHPO
014064;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=972;
                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                               SCHPO
                                                                                                                                                                                                                                                                                                          1250
                                                                 321
                              a
                                                                                                 g
                                                                                                                                                                 a
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                           ò
                                                               ò
                                                                                                                               òγ
                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licensee agreement (see http://www.isb-sib.ch/announce/or send an email to licenseelab-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       analysis of the dealers of 1994).

J. Bacteriol. 176:389-3850(1994).

-!- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EXTRACELLULAR (BY SIMILARITY).
MEMBRANE ANCHOR (BY SIMILARITY).
MEMBRANE (BY SIMILARITY).
CONSERVED IN GRAM-POSITIVE COCCI SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -SLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEP-DSPTTTDPD-----AAASAT 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           681 TDAKGITGNOVWTYGKKGDNFRTVQLLNLMGINSDWKNEDGSAANKTPDEQTNLTVKYAL 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDVSMEDBACRAVITYVTSPDDWSKSNLQKVSASVKTDENGKPVLVINVPKLTLMDVVYI 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SNETPATĞĞNSEQPELN---EPTAQTQPSSQVSPANTSVTPVAEQPTNQGQAADKADQAP 977
                                                                                                                                                                               PELLICLE-COATED TOOTH SURFACE. THE ACTIVITY OF THIS ENZYME IS OPTIMAL AT PH 5.3 AND AT 39 DEGREES CELSIUS.
-!- CATALYTY: ENDOHYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC LINKAGES IN DEXTRAN.
-!- SUBGUNIT: HOMODIMER.
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
-!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS IN THE REGION OF THE MEMBRANE ANCHOR.
-!- SIMILARITY: BELONGS TO FAMILY 66 OF GLYCOSYL HYDROLASES.
-!- SIMILARITY: BELONGS TO FAMILY 66 OF GLYCOSYL HYDROLASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEITLNDGFEVLHD---HGPLDTLNRQIGSSVFRVETQEDGKHIAV-------121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---GQRNGVET-----SVVLSDQEYARLQSIDPEGKDKFVFT-------GGRGGA- 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                801 SNANQESAPEADQAQTPAAQSSDDKVAENETSQPAAEDAKEQTSEPAQDQAAPAEQGQAI 860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.5%; Score 127; DB 1; Length 1337;
19.0%; Pred. No. 6.3;
tive 88; Mismatches 248; Indels 230; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGD-----NRASDVPGLPVN-PMRLAA 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 -----GHAMVTVAS----DITEARQRILELLEPKGTGESKGAGES-----
                                                                                                 of Streptococcus sobrinus
Escherichia coll and sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
Hydrolase; Glycosidase; Slgnal; Transmembrane; Repeat; Plasmld.
SIGNAL 1 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MW; B494275A77A2E3D0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----KGVGELRESNSGAENTTETQTSTSTS----
                               FROM N.A., AND SEQUENCE OF 31-36.
15 / UAB66;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEXTRANASE
                                                                                               "Purification and characterization dextranase produced in recombinant analysis of the dextranase gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00746; Gram_pos_anchor; 1.
PROSITE; PS00343; GRAM_POS_ANCHORI
                                                               MEDLINE-94292401; PubMed-8021165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M96978; AAA21772.1; -.
                                                                                   Wanda S.-Y., Curtiss R. III;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1337 AA; 143298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR001899; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1314
1333
1305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 133;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
DOMAIN
                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   921
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                        1036 ETPVADOVPAVAEQPQATEPNOAKPSVDKAAAPEALSLIOLKOOTPAIQAKEADDPEVDE 1095
                                                                                                                               1096 TKSEVTPDSGTDKAPEAGQVDSDKAPTVKPSTPENNDNQPNNANDADKNKTNEADSNKAN 1155
                                                                                                                                                                                                                                                                                                                                                                                             Uren A.G., Beilharz T., O'Connell M.J., Bugg S.J., van Driel R., Vaux D.L., Lithgow T., "Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell division.";
ETATRDQLTKEAFQ----NPDNQKVNID------ELGNAIPSGVLKDDVVANIEE
                                                                                      ------QAKAGE-----EAKQQAIENN-----AQAQKKYDEQQAKRQ
                                                                                                                                                                                                     353 EELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVENKP
                                                                                                                                                                                                                                                                413 A----NNT---PAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLMLKH
                                                                                                                                                                                                                                                                                                                                                  465 RCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPS----AGIQSTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Proc. Natl. Acad. Sci. U.S.A. 96:10170-10175(1999).
-!- FUNCTION: SEEMS TO ACT IN THE PLEIOTROPIC CONTROL OF CELL
DIVISION. MAY PARTICIPATE IN CHROMOSOME SEGREGATION EVENTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Morishita J., Matsusaka T., Yanagida M.; "Fission yeast cutl7 is required for chromosome segregation."; Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Harris D., Wood V., Rajandream M.A., Barrell B.G.;
Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-UUL-1998 (Rel. 36, Created)
15-UUL-1998 (Rel. 36, Last sequence update)
01-CCT-2000 (Rel. 40, Last annotation update)
BIRI PROTEIN (CHROMOSOME SEGREGATION PROTEIN CUT17).
BIRI OR CUT17 OR SPCC962.02C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1282 --VATSPEKKSEPVSKT----STTSSSDKLPKTGDHKTV 1314
                                                                                                                                                                                                                                                                                                                                                                                                                                          521 ARLALSGGLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -! - SIMILARITY: CONTAINS 2 BIR REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-99398681; PubMed-10468581;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AB031034; BAA83415.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAA20434.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001370; -. Pfam; PF00653; BIR; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Schizosaccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-4896;
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          339 Q---AQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTT 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                396 TITITITISARIVENKPANNTPA-----QGNVDTP-----GSEDIMES---RRS 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                223 SSL----RSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETA 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      279 TRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNA 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            568 SQTAIHVSKFEDLENKSMESEQSLQLLSESENDDKPLIDLIPLLAIKRKDNLVSGVLEKG 627
                                                                                                                                                                                                                                                                                                                                       73 PMRLAASEITLINDGFEVLHDHGPLDTLNRQIGSS------VFRVETQED----- 115
                                                                                                                                                                                                                                                                                                                                                                                                            116 -----GKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTV 168
                                                                                                                                                                                                                                                                                                                                                                                                                                       325 PFSKGICNDSMQVAKKNFTEEIPLKEDEKDNELEHLVSPATS--VHTTVSDITGHQSVTD 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              169 ASDITEARORILELLEPKGTGESKGAGESKGVGELRESNSGA-----ENTTETQTSTST 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    439 PEQQKVEKEDEHLNL-----MGSFIEESTKQPISSKPSTSSPD----MTDAA 481
                                                                                                                                                                                                                                                                                           222 DNSILPVSPTRDSTKSHKTLNFSPSRKNNLNARPLIMSLYTNTSEEKOSQPTRAPQSPTK 281
                                                                                                                                                                                                                                                                                                                                                              1: | | : : | | : : | | 324 PVLLTAPR------RKNKSPKKSPKPKPAVFKPVKPIFSDEDEDDDLTASQ 324
                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                   13 NNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNRASDVFGLPVN 72
                                                                                                                                                                                             Query Match

4.4%; Score 126; DB 1; Length 997;
Best Local Similarity 19.8%; Pred. No. 5;
Matches 98; Conservative 62; Mismatches 205; Indels 130;
PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG.
PROSITE; PS50143; BIR_REPEAT_2; 2.
Cell division; Repeat.
REPEAT 120 194 BIR 1.
REPEAT 120 194 BIR 2.
DOMAIN 80 83 POLY ASP.
DOMAIN 487 490 POLY SER.
SEQUENCE 997 AA; 112579 MW; 952A6BAFA5C489F4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          437 SMASTSSTFFDTSSI 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||| | |||| :
628 KSTSTSKTKFDTSIV 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                 ð
            DR DR DR ETT FT FT SO
```

Search completed: September 27, 2001, 14:22:56 Job time: 152 sec

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                        09R396
                                                                                                                                                                                                                                                                                       RESULT
Q9R396
                                                                                                                                                                                                                                                                                                           ; Search time 38.06 Seconds (without alignments)
1943.208 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9r396 escherichia
085506 escherichia
047014 escherichia
047016 escherichia
085508 escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  drosophila
caenorhabdi
caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             escherichia
escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9vtk8 drosophila
Q9n973 leishmania
Q9vkjl drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                citrobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              drosophila
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                             1 MPIGNLGHNPNVNNSIPPAP.....SNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          068258
09kwh9
052147
09wxk1
09etil
050190
09nhx4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q94185
P91365
O77261
Q9w5e0
                                                                                                                                                                               425026
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                           425026 seqs, 132305027 residues
                                                       September 27, 2001, 14:27:03
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                   summaries
                                     OM protein – protein search, using sw model
                                                                                                                                         Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09R396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       047016
085508
068258
09KWH9
052147
09WXK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        050190
09NHX4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          094185
P91365
O77261
Q9W5E0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9VTK8
Q9N973
Q9VKJ1
                                                                                                                                                                                                                                                                                                                                                                                   sp_unclassified:*
                                                                                                                                                                                                                                                                                              sp_fung1:*
sp_human:*
sp_invertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                            sp_vertebrate:*
                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                  SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                             sp_organelle:*
                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                           US-09-189-415A-4
2851
                                                                                                                                                                                                                                                                                                                                                                sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                             sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                     sp_virus:*
                                                                                                                                                                                                                                                                                                                                                      sp_phage:*
                                                                                                                                                                                                                                                                                                                                     sb_mhc:*
                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          555.1
54.3
53.4
53.4
                                                                                                                                                                                                                                                                                                                                                        ......
1122....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2608.5
1774.5
1766.5
1753.5
1554.5
1554.5
1522
1522
                                                                                                     Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           170
165.5
162.5
162.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
156
152.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   190
                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2624.
                                                                                                                                                                                                                                                                     ..
                                                                                                               Sequence:
                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                    Database
                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . 8
1110
1110
1110
1110
1110
1110
                                                                                            Title:
```

Q9vb94 drosophila	Q9p319 neurospora	Q17346 caenorhabdi	Q9tw45 caenorhabdi	097205 leishmania	Q9v8q1 drosophila	Q23587 caenorhabdi	077509 bos taurus	076602 caenorhabd1	Q63661 rattus norv	Q9nds4 dictyosteli	Q9v8b9 drosophila	Q9w3u2 drosophila	Q17368 caenorhabdi	Q9tvg6 caenorhabd1	Q9ptk7 xenopus lae	Q9gyw7 drosophila	Q9w468 drosophila	Q22248 caenorhabdi	Q9kjj3 enterococcu	Q912m3 pseudomonas	091d52 oryza sativ	Q9w4z0 drosophila	Q9w2z3 drosophila	Q9vk10 drosophila	P90666 anthocidari	
Q9VB94	Q9P3I9	017346	Q9TW45	097205	097801	023587	077509	076602	063661	09NDS4	Q9V8B9	Q9W3U2	017368	Q9TVG6	Q9PTK7	Q9GYW7	Q9W468	022248	Q9KJJ3	Q912M3	Q9LD52	Q9W4Z0	Q9W2Z3	Q9VKL0	P90666	
Ŋ	m	'n	Ŋ	ഹ	Ŋ	Ŋ	9	ស	11	ഗ	Ŋ	ഹ	S	ស	13	S	ស	Ŋ	7	7	10	S	Ŋ	വ	5	
949	770	1192	1192	3012	332	3507	1323	1275	1805	2678	1026	1086	1096	1144	606	1475	875	1829				1385	4498	620	837	
5.2	5.2	5.2	5.2	5.2	5.1	5.1	5.1	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.9	4.9	4.9	4.9	4.9	4.9	4.8	4.8	4.8	4.8	4.8	
149.5	148.5	148	148	147.5	145	144.5	144	143.5	143.5	143	142.5	141.5	141.5	141.5	141	141	140.5	140.5	140	139	138	137.5	137.5	137	137	
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	

ALIGNMENTS

```
STRAIN=EDL933;
MEDLINE=98339885; PubMed=9673266;
MEDLINE=98339885; PubMed=9673266;
Perna N.T., Maybew G.F., Posfai G., Elliott S., Donnenberg M.S.,
Kaper J.B., Blattner F.R.;
"Molecular evolution of a pathogenicity island from enterohemorrhagic Escherichia coli 0157:H7.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       is
                                                                                                                                                                                         Escherichia coli.
Bacteria; proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                    Finlay B.B.;
"Enterohemorrhagic Escherichia coli 0157:H7 produces Tir, which translocated to the host cell membrane but is not tyrosine phosphorylated.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 558;
                                                                                                                                                                                                                                                                                                                                                        STRAIN-86/24;
MEDLINE-99442825; PubMed-10225900;
DeVinney R., Stein M., Reinscheid D., Abe A., Ruschkowski S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;
                                                   01-WAY-2000 (TrEMBLrel. 13, Created)
01-WAY-2000 (TrEMBLrel. 13, Last sequence update)
01-WAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 2624.5; DB 2;
Pred. No. 1e-152;
  558 AA.
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Infect. Immun. 66:3810-3817(1998)
EMBL, AF125993; AAD29391.; --
EMBL, AF071034; AAC31506.1; --
InterPro; IPR003536; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nfect. immun. 67:2389-2398(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PR01370; TRNSINTIMINR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92.1%;
  PRELIMINARY;
```

ď

a

3;

Gaps

7;

Indels

27;

Mismatches

3;

Conservative

Matches 525;

3;

Gaps

7:

Indels

24;

3; Mismatches

Conservative

528;

g

a à g

ò

Q δ g à g ð Ω ò g ð 셤 ŏ g

οŽ

9

360

360

240

300 300 420 420 479

537

477

```
VGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
                                                                  NRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-98294040; PubMed-9632251;
MEDLINE-98294040; PubMed-9632251;
Deibel C., Kraemer S., Chakraborty T., Ebel F.;
"Espe, a novel secreted protein of attaching and effacing bacteria, directly translocated into infected host cells, where it appears as tyrosine-phosphorylated 90 kDa protein.";
Mol. Microbiol. 28:463-474(1998).
                              LGNA I PSGVLKDDVVAN I EEQAKAAGEEAKQQA I ENNAQAQKKY DEQQAKRQEELKVSSG
                                                                                                                                                                                                                                                                                             AGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPANNTPAQG
                                                                                                                                                                                                                                                                                                                                                 GLIGLAATGIVQALALIPEPDSPITIDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE
                                                                                                                                                                                                                                                                                                                                                                                      NVDTPGSEDTMESRRSSMASTSSTFFDTSSIG---GPCRIRMLMLKHRCMIRRCRLLILI
                MPIGNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Krejany E.O.; Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=REPEC 84/110/1, AND E65/56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 047014 PRELIMINARY;
047014;
01-00V-1996 (TTEMBLEEL 01,
01-MAY-1999 (TTEMBLEEL 10,
01-MAR-2001 (TTEMBLEEL 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          537 GGSNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGSNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-95ZG1;
                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               538
                                                                                                                                                                                                                                                                                                                                                                                                                   421
                                                                                                                                                    121
                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                          421
                                                                                                                                                                            181
                                                                                                                                                                                                                                 241
                                                                                                61
                                                                                                                          121
                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                      ò
                                            q
                                                                        ò
                                                                                               g
                                                                                                                          δ
                                                                                                                                                 g
                                                                                                                                                                            οy
                                                                                                                                                                                                    q
                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                             δŽ
                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       536
                                                                                                                          240
                                                                                                                                                                                                         240
                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                                                                      360
                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                            477
                                                                                                                                                                                                                                                                                                                                                                                                                    479
                                                                                                                                                                                                                                                                                                                                                                                                                                             537
                                                                        NRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
                                                                                      300
                      9
                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                           480 QNM---GNTDSVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLT
                                   GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE
                                                                                                                                                                                                                                                                                                     MPIGNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGD
                                                                                                                                                                                ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETOTSTSSIRSDPKLMLALGTVAT
                                                                                                                                                                                                                                   GLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDE
                                                                                                                                                                                                                                                                                      LGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSSG
                                                                                                                                                                                                                                                                                                                                         AGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPANNTPAQG
                                                                                                                                                                                                                                                                                                                                                                                             421 NVDTPGSEDTMESRRSSMASTSSTFFDTSSIG---GPCRIRMLMLKHRCMIRRCRLLLLL
                                                                                                                                                                                                                                                                                                                                                                                                                                              RLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
The translocated intimin receptors (Tir) of Shiga toxigenic
Escherichia coli isolates belonging to serogroups 026, 0111, and
react with sera from patients with hemolytic-uremic syndrome and
eschibit marked sequence heterogeneity.";
Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL, AF070067; AAC69314.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 558;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CA2CDDAC94527C2E CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 2608.5; DB 2
Pred. No. 9.5e-152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 538 GGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRNSINTIMINR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91.5%;
93.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 08, (TrEMBLrel. 08, (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1998 (TrEMBLrel.
01-NOV-1998 (TrEMBLrel.
01-MAR-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR003536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   558 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PR01370;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Escherichia coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Escherichia.
NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-95SF2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         085506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           085506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N
                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                478
                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                                      421
                        Н
                                                                             61
                                                                                                    61
                                                                                                                                                                                 181
                                                                                                                                                                                                                                    241
                                                                                                                                                                                                                                                              241
                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                           361
Matches
```

RESULT 085506

is a

 \sim

```
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PR01370; TRNSINTIMINR
                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, U59504, AAD19750.1; -. EMBL, AF045568, AAC15683.1; -EMBL, U59503, AAB02941.1; -. InterPro, IPR003536; -.
                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 172-538 FROM N.A
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    359; Conservative
                                                                                                                                                                                              O'Brien R.A.;
Submitted (DEC-1996)
                                                                                                                                                                                                                                                                                   Submitted (MAR-1999)
                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-REPEC RDEC-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                     SEQUENCE FROM N.A. STRAIN-REPEC 83/39;
                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN=REPEC 83/39.
                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                               Escherichia coli
                                                                                                                                         NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                              STRAIN-RDEC-1;
                                                                                                                                                                                                                                                                    Krejany E.O.;
                                                                                                                                                                                                                                                                                                                                                                                                                   Krejany E.O.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                    .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHI 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 AVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLMLALGTVA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSS 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPANNTPAQ 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GNVDTPGSEDTMESRRSSMASTSSTFFDTSS-----IGGPCRIRMLMLKHRCMIR 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          470 RCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGL 529
          Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
"Translocated intimin receptors (Tir) of Shiga-toxigenic Escherichia coli isolates belonging to serogroups 026, 0111, and 0157 react with sera from patients with hemolytic-uremic syndrome and exhibit marked sequence heterogeneity."
Infect. Immun. 66:5580-5586(1998).
                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                              1 MPIGNLGHNPNVNNSIPPAAPPLPSQTDGAGG-RGQLINSTGPLGSRALFTPVRNSMADSG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                43;
                                                                                                                                          3
                                                                                                                                                                                                                                                                                                                                                                     62.2%; Score 1774.5; DB 2; Length 538; 63.2%; Pred. No. 8.5e-101; ive 61; Mismatches 106; Indels 43;
                                                                                                                                       Marches O., Nougayrede J.-P., Boury M., Mainil J., Charlier Boullier S., De Rycke J., Milon A., Oswald E., "Role of Tir and Intimin in the pathogenesis of rabbit enteropathogenic Escherichia coll."
                                                                                                                                                                                                Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases EMBL; U59502; AAC32028.2; -.
                                                                                                                                                                                                                                                                                Interpro; IPR003536; -...
PRINTS; PR01370; TRNSINTIMINR.
Hypothetical protein; Receptor.
SEQUENCE 538 AA; 55420 WW; 31D7A8E227B3D06C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  530 RHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-99003184; PubMed-9784578;
                                                                                                                                                                                                                          EMBL, AJ223063; CAA11065.1; -.
EMBL, ARC70068; AAC6316.1; -.
EMBL, AF112728; AAD27866.1; -.
EMBL, AF113597; AAF03080.1; -.
                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 63.23
Matches 360; Conservative
                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-B10;
                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 420 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               509
```

ð a ò g ò g à q ŏ a ŏ g à g ò g δ g ò g

```
60 DNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHI 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 IDKLEPKDTKETKEPG------DPNSGGKIIEIHTSTSTSSLRADPKLWLSLGTIA 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         300 ELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSS 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360 GAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSARTVENKPANNTPAQ 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        352 GVGYGISGALILGGGIGAGVTAALHRKNQPAEQTITTRT------VVDNQPTNNASAQ 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             420 GNVDTPGSEDTMESRRSSMASTSSTFFDTSS-----IGGPCRIRMLMLKHRCMIR 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGAGG-RGQLINSTGPLGSRALFTPVRNSMADSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                240 TGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNID
                                                                                                                                                    Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62.0%; Score 1766.5; DB 2; Length 63.0%; Pred. No. 2.6e-100; ive 61; Mismatches 107; Indels
                                                                                                                                                                                                                                                                                                                                               to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Agin T.S., Boedeker E.C.;
Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           538 AA; 55410 MW; 3132A969B7B3D06C CRC64;
                      01-AUG-1999 (TrEMBLrel. 11, Last sequence update) 01-MAR-2001 (TrEMBLrel. 16, Last annotation update) TRANSLOCATED INTIMIN RECEPTOR.
01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
NCBI_TaxID=562;
                                                                                                                                                                                                            Escherichia
                                                                                                                                                                                                                                                                                        syndrome."
                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                 068258
                                                                                                                                           068258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           352
                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292
      104
                                            464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                        οy
                                           a
                                                              ŏ
                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRSDPKLWLALGTVA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240 TGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNPDNQKVNID 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKRQEELKVSS 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHI 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0157
                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                            GAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVENKPANNTPAQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNVDTPGSEDTMESRRSSMASTSSTFFDTSS-----IGGPCRIRMLMLKHRCMIR
470 RCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGL
                                                                                                                                                                                                                                                                       Parton M. Manning P.A., Woodrow M.C., Paton J.C.;
Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
"The translocated intimin receptors (Tir) of Shiga toxigenic
Escherichia coli isolates belonging to serogroups 056, 0111, and
react with sera from patients with hemolytic uremic syndrome and
exhibit marked sequence heterogeneity.";
Submitted (JUN-1998) to the EMBL, GenBank/DDBJ databases.
EMBL; AF070069; AAC69318.1; ...
InterPro; IPR03785; ...
PRINTS; PR01370; TRNSINFIMINR.
SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;
                                                                                                                                                                                                              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                            43;
                                                                                                                                                                                                                                                                                                                                                                                          2; Length 538;
                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                     01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-NAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                         61.5%; Score 1753.5; DB 2; 62.6%; Pred. No. 1.6e-99; ive 61; Mismatches 109;
                                                                                                                                      Ź
                                                                                                                                      538
                                                                             530 RHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 62.6
Matches 357; Conservative
                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                         Escherichia coli
                                                                                                                                                                                                                                    NCBI_TaxID=562;
                                                                                                                                                                                                                                                                  STRAIN-EPEC87A;
                                                 E----E
                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                       085508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                  464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             φ
                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                              셤
                                                                    ò
                                                                                     g
                              à
```

```
412 PANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFF-DTSSIGGPCRIRMLMLKHRCMIRR 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSARTVENK 411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNOKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQAKR 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      117 KHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QRILELLEPKGTGESKGAG---ESKGVGELRESNSGAENT--TETQTSTSTSSLRSDPKL 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59 GDNRASDVPGLPVNPMRL--AASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDG 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                   1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
Voss E., Paton A.W., Manning P.A., Paton J.C.;
"Molecular analysis of Shiga toxigenic Escherichia coli Olll:H-
proteins which react with sera from patients with hemolytic-uremic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  470 RCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGL
                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6EC95F76EF0F44CC CRC64;
                                                                                                                                                                                                                                                                                      01-AUG-1998 (TrEMBLrel. 07, Last sequence update) 01-MAR-2001 (TrEMBLrel. 16, Last annotation update) TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67; Mismatches 134;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 55.1%; Score 1570; DB 2; Best Local Similarity 59.5%; Pred. No. 2.7e-88; Matches 339; Conservative 67; Mismatches 134;
                                                                                                                                                                                                                                                        Ŕ
                                                                                                                                                                                                                                                      551
                                                                                                                            530 RHDMGGLTGGSNSAVNTSNNPPAPGSHRFV
                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE=98187918; Pubmed=9529069;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Infect. Immun. 66:1467-1472(1998)
EMBL; AF025311; AAC69249:1; --
PRINES; PR01370; TRNSINTIMINR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56975 MW;
                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       551 AA;
```

```
SEQUENCE FROM N.A.
Biliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
Elliott McNamarra B.P., Donnenberg M.S., Kaper J.B.;
Mol. Microbiol. 0.0-0(1998).
EMBL; AF022236; AAC38390.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  59 GDNRASDVPGLPVNPMRLAA--SEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDG 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117 KHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEA 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FONPDNOKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQ 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VENKPANNTPAQCNVDTPGSEDTMESRRSSMASTSSTFFDTSS------IGGPCRIR 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MLMLKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPP-----RDTTDNG---ARLLG 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VOQQTGGNTPAGGGTDATRAEDASLNRRDSQGSVASTHWSDSSSEVVNPYAEVGG---AR 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          462 NSLSAHQ--------PEEHIYDEVAADPGYSVIQNFSGSGPVTGRLIG 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          177 QRILLELLEPKGTG-----ESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSART
                                                                                                                                                                                                                                             Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     511 NPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                                                                                                                                                                                                                                                                                                                                                                                                                                    54.3%; Score 1540.5; DB 2; Length
56.0%; Pred. No. 5.6e-87;
.ive 67; Mismatches 123; Indels
                                                                                                                                                                                                                                                                                                                                                                                             19DD08A9BE9251CB CRC64;
                                                                                                                                                            01-MAR-2001 (TIEMBLrel. 06, Last sequence update)
TRANSLOCATED INTIMIN RECEPTOR.
TIR.
                                                               509 STYALLANSAGLRIGMGGLTGSGESAVNTANANAAPTPGPVRFV
                                       518 STYARLALSGGLRHDMGGLTGGSNSAVNT--SNNPPAPGSHRFV
                                                                                                                                   550
                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                               PR01370; TRNSINTIMINR.
5 550 AA; 56509 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR003536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                Escherichia coli
                                                                                                                                                                                                                                                                         NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Sim
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Si
Matches 330,
                                                                                                                                                                                                                                                                                                                                                                                                SEOUENCE
                                                                                                                                                                                                                                                                                                                                                                                  PRINTS;
                                                                                                                                 052147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    405
            461
                                                                                                       RESULT
                                                                                                                      052147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
            a
                                      δ
                                                              g
                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDNRASDVPGLPVNPMRLAA--SEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDG 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    117 KHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FONPDNOKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQ 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSS------IGGPCRIR 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               469 AEEHIYDE-----VAADPNYSVIQNFSGNNQVTG-RLMGTPGQGIQSTYAILINNSAGL 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSART 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MLMLKHRCMIRRCRLLILIRLFRIWG-IQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQ 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MPIGNLGNNVNSNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLFSPLRNSVADS 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MPIGNLGHNPNVNNSIPPAAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS 58
                                                                                                                                                                                                                                                                                                                                                                  "Analyses of type III secreted proteins and Tir in enteropathogenic Escherichia coli 0157:H45.";
Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB036053; BAA96815.1; ---
InterPro; IPR003536; ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QRILELLEPKGTG------ESKGAGESKGVGELRESNSGAENTTETQTSTSTSSLRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEA
            471 CRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLA-LSGGL
                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         552;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 54.5%; Score 1554.5; DB 2; Length Best Local Similarity 57.5%; Pred. No. 2.4e-87; Matches 336; Conservative 63; Mismatches 128; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                 57005 MW; A8D79EE22EE50A4B CRC64;
                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                                                                                                                                                              Ä.
                                                                                                                                                              552
                                                                               RHDMGGLTGGSNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                      PR01370; TRNSINTIMINR
                                                                                                                                                                                       15,
15,
16,
                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                      01-OCT-2000 (TrEMBLrel.
01-OCT-2000 (TrEMBLrel.
01-MAR-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   552 AA;
                                                                                                                                                                                                                                                                                                                                                     Abe A., Nagano H.;
                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                           co11,
                                                                                                                                                                                                                                                                                           NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                          STRAIN-HK01;
                                                                                                                                                                                                                                                         Escherichia
                                                                                                                                                                                                                                                                                    Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; P
SEQUENCE
                                                                                                                                                                            09КМН9;
                                                                                                                                                            09KWH9
                                                                530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     459
                                                                                                                                 RESULT
Q9KWH9
                                    g
                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                               ò
```

Length

407

227

g

```
Citrobacter
                                                                                                                                                                                                                                                                                                                    Receptor.
SEQUENCE
                                                                         Q9ETI1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   468
                                                     RESULT 10
                                                                 Q9ETI1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
ò
                   g
                                                                                     10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTSARTVE 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NKPANNTPAQGNVDTPGSEDTMESRRSSMASTSST-FFDTS-SIGGPCRIRMLMLKHRCM 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              404 QQTGGNTRAQGGADTTGVENASLTRRDSQASVASTQWSDTSGDVVNP-----YAEGWM 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IRRCRLLILIRIFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSG 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQR 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKH 118
                                                                                                                                                                                                                                                                                                                                                                                        38; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                        Okutani A., Itoh K., Sasakawa C.;
"Translocated intimin receptor(Tir) of murine pathogenic Escherichia coli 0115a.c:K(B).";
Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
EMBL. ABU26719; BAA77400.1; -.
InterPro; IPR003484; -.
                                                                                                                                                                                                                                                                                                                                                                                                              1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                        179 ILELLEPKGTG------ESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDP
                                                                                                                                                        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                   Length 547;
 Indels
                                                                                                                                                                                                                                                                                                                                    02CAC6D625FA6EE1 CRC64;
                                                                                        01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-NAR-2001 (TrEMBLrel. 16, Last annotation update)
TRANSLOCATED INTIMIN RECEPTOR.
                                                                                                                                                                                                                                                                                                                                                                  Query Match 53.4%; Score 1522; DB 2; Best Local Similarity 57.2%; Pred. No. 2.3e-85; Matches 327; Conservative 68; Mismatches 139;
                                                                    Ş
                                                                    547
                                                                    PRT;
                                                                                                                                                                                                                                                                                                   PRINTS; PR01370; TRNSINTIMINR.
ProDom; PD000551; -; 1.
                                                                                                                                                                                                                                                                                                                                    56283 MW;
                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                    547 AA;
                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                Escherichia coli
                                                                                                                                                                    Escherichia.
NCBI_TaxID=562;
                                                                                                                                                                                                                 STRAIN-MPEC;
                                                                                                                                                                                                                                                                                                                           Receptor.
                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  350 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              410
                                                                    Q9WXK1
                                                                                Q9WXK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            291
                                              σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             457
                                                        Q9WXK1
```

g ò g

A STANDAR STAN

g ò g δ 셤 à qq õ g δ g

임 ð

ò

```
SEQUENCE FROM N.A.
STRAIN-CDC 1843-73T, AND DBS100;
STRAIN-CDC 1843-73T, AND DBS100;
STRAIN-CDC 1843-73T, Dand-1101562;
Luperchio S.A., Newman J.V., Dangler C.A., Schrenzel M.D.,
Brenner D.J., Steigerwalt A.G., Schauer D.B.,
"Citrobacter rodentium, the Causative Agent of Transmissible Murine
Colonic Hyperplasia, Exhibits Clonality: Synonymy of C. rodentium and
Mouse-Pathogenic Escherichia coli.";
J. Clin Microbiol. 38:434-4350(2000).
EMBL; AF301618; AG40788-1;
EMBL; AF301617; AAG25642.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 IAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQR 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         350 KRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTTTSARTVE 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              351 KRQQELDLSSGIGYGLSSALIVGGGIGAGVTAMLHRRNPPTEQTIATTHS-----VIQ 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            410 NKPANNTPAQGNVDTPGSEDTMESRRSSMASTSST-FFDTS-SIGGPCRIRMLMLKHRCM 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IRRCRLLILIRLFRIWGIQISVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MPIGNLGHNPNVNNSIPPAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59 GDNRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     230 KLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEAFQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    179 ILELLEPKGTG------ESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDP
                                                                                                                                                                                                                                                                              Citrobacter rodentium.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 547;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 547 AA; 56270 MW; CB8318B301049C37 CRC64;
                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53.3%; Score 1521; DB 2;
57.2%; Pred. No. 2.7e-85;
Live 68; Mismatches 139;
                                                                                                                                                547 AA
559
                                     547
528 GLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV
                      Created)
                                                                                                                                                PRT;
                                                                                                                                                                                     01-MAR-2001 (TrEMBLrel. 16, Create
01-MAR-2001 (TrEMBLrel. 16, Last s
01-MAR-2001 (TrEMBLrel. 16, Last s
TRANSI,OCATED INTIMIN RECEPTOR TIR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 53.3'
Best Local Similarity 57.2
Matches 327; Conservative
                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=67825;
```

```
01-0CT-2000 (TrEMBLrel. 15, 01-0CT-2000 (TrEMBLrel. 15, 01-MAR-2001 (TrEMBLrel. 16, ALHAMBRA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                        PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00249; PHD; 1
                                                                                                                                                                                                                                                                                       ALHAMBRA OR CG1070.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 131; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                           Q9NHX4;
                                                                                                                                                      Q9NHX4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             691
                                                                                                              RESULT
                                                                                                                                   7XHN60
                                       qq
                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDNRASDVPGLPVNPMRLAA - - SEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDG 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            179 TKILAKLDPDNHGGRQPKDVDTRSVGVGSASGI-----DDGV--VSETHTSTTNSSVRS 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DPKFWVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDQAANAAESATKDQLTQEA 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VENKPANNTPAQGNVDTPGSEDTMESRRSSMASTSSTFFDTSS------IGGPCRIR 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VVQQQTGGIPQHKVALMPQERRRFSDRRDSQGSVASTHWSDSSSEVVNPYAEVGG---AR 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MLMLKHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHPP----RDTTDNG---ARLLG 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DPKLWLALGIVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQLTKEA 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTTTTTTTTTSART 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Enteropathogenic E. coli (EPEC) transfers its receptor for intimate adherence into mammalian cells.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MPIGNLGHNPNVNNSIPPAAPPLPSQTDGA--GGRGQLINSTGPLGSRALFTPVRNSMADS 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MPIGNLGNNVNGNHLIPPAPPLPSQTDGAARGGTGHLISSTGALGSRSLESPLRNSMADS 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QRILELLEPKGTG-----ESKGAGESKGVGELRESNSGAENTTETQTSTSTSLRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FONPDNOKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIENNAQAQKKYDEQ
                                                                                                                                                                                                                                                                                                                        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 549;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kenny B., DeVinney R., Stein M., Reinscheid D.J., Frey E.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       549 AA; 56843 MW; 40C8B8B234409A08 CRC64;
                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 52.4%; Score 1495; DB 2;
Best Local Similarity 54.7%; Pred. No. 1e-83;
Matches 322; Conservative 65; Mismatches 132;
GLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV 559
                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=E2348/69;
MEDLINE-98050926; PubMed=9390560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; AF013122; AAB88410.1; -. InterPro; IPR003536; -. PRINTS; PR01370; TRNSINTIMINR.
                                                                                                                                                                                             01-JUN-1998 (TrEMBLrel. 06, C)
01-JUN-1998 (TrEMBLrel. 06, Ld
01-MRR-2001 (TrEMBLrel. 16, Lk
TRANSLOCATED INTIMIN RECEPTOR
                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cell 91:511-520(1997
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                              001
                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Finlay B.B.;
                                                                                                                                                                                                                                                                                                          Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                           050190;
                                                                                                                                                 050190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               461
                                                                                                     RESULT
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
050190
05
                                    g
  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
```

```
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryyota; Neottera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBL_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ASDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIAVG 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---F 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      416 SANNNFTNSSASSTSSNSSSTREKSSSKLSKNKDSNQVPSATSSLSTTSSINTQPSSSTS 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153 VFTGGRGGAG-HAMVTVASDITEARQRILE-----LLEPKGTGESKGAGESKGVGELRE 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SNSG-----AENTTETQTSTSTSSLRSDPKLWLALGTVATGLIGLAA--TGIVQ 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          253 ALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKD 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --EQQAKRQE 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----SAKSTSSISTSGSALEDNNSLISRYDIKDVQVALTPLTDFEKEIEKSSKRQR 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -ELKV---SSGAGYGLSGALILGGGIGVAVTAALHRKNOPVEQTTTTTTTTTTTTARTVE 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NKPANNTPAQGNVDT-----PGSEDTMESRRSSMASTSSTFFDTSSIGGPCRIRMLML 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGNSSGVSATGGAQSAVGSGGYPKTESSKSSGTASAGSGSGSSSNTSSTKHGSNIKDISS 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        463 KHRCMIRRCRLLILIRLFRIWGIQISVVYSTIQHP----PRDTTDNG----ARLLGNPSA 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IGNLGHNPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSMADSGDNR 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             536 SSSGFGSDLRSVSTSSSSTVNDSTGGFGSNSNSERENLSGAGSSASNMPGTIAPGTGGVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF-17.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.7%; Score 190; DB 5; Length 1323;
20.2%; Pred. No. 0.0019;
ive 90; Mismatches 273; Indels 154;
                                                    549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.

Perrin L., Dura J.M.;

Alhambra, a Drosophila homolog of mammalian AF-10 and A Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.

EMBL, AF217960; AAF72595.1;

FlyBase; FB900037471; Alhambra.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   D53C0C8AF392F9A6 CRC64;
                        511 NPSAGIQSTYARLALSGGLRHDMGGLTGGSNSAVNTSNNPPAPGSHRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SSAATNLSTNKGGSSSSTANSLTSTST----SSGSSSNSSSKKRKAD-
                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123 QRNGVET----SVVLSDQEYARLQSIDPEGKDK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               313 DVVANIEEQAKAAGE-EAKQQAIENNAQAQKKYD---
                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1323 AA; 132883 MW;
```

```
VARSPLIC
                                                                                                                                                                                                                                                                                                                                             O.F.
                                                                      P91365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1650
                                          RESULT 14
                                                         P91365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYSTEKTSTIKKASTIEEPTITIESSTICKATIPELSTIKEETTIEKITIEG 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         333 AIENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQ 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQ 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIKEAF------QNPDNQKVNIDELGNAIPSGVLKDDV-VANIEEQAKAAGEEAKQQ 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            705 ITQTSVSVVESSTPRQLPERWKAIVNKFKHNLE--VLKEKKRLLKEKESTSTTGSDSS-- 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TVASDITEARQRILELLEPKGTGESKGAGESKGVGELR----ESNSGAENTTETQTSTST 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STTTEEP-----TTTAIFAEASTGII----TTDEETTSTTSTTPEITSTKEIVTESA 704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --ETTTVVAENIDEVITTEKEKV---------VOTTPITTEKSTTOEE 797
811 SSNQQASTASSSSAPSLY----VSVPLSTANVPGINLPTSSTSSSTTSESHSASSRSS 864
                                                                                                                                                                                                                                                                                                                                                            ., Berks M.,
Cooper J., Coulson A.,
                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITITITITITARIVENKPANNIPAQGNVDIPGSEDTMESRRSSMASISSIFFIIS 450
                                                                                                                                                                                                                                                                                                                                                         Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bunfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Jones M., Green P., Hawkins T., Hiller L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
Barsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen F.
Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
Thierry Mieg J., Thomas K., Yaudin M., Vaudin M., Walson A., Welldman P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6.0%; Score 170; DB 5; Length 1229; 22.8%; Pred. No. 0.029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44; Mismatches 122; Indels
                               515 GIQSTYARLALSGGLRHDMGGLTG------GSNSAVNTSNNPPAP 553
                                                EMBL; U67956; AAB07691.1; -.
SEQUENCE 1229 AA; 133101 MW; 7761C45FFCBE937D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                           01-FEB-1997 (TrEMBLrel. 02, Last sequence update) 01-NOV-1998 (TrEMBLrel. 08, Last annotation update) 000DE FOR BY C. ELEGANS CDNA YK126F9.5.
                                                                                                                                  PRT; 1229 AA
                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                STRAIN-BRISTOL N2;
MEDLINE-94150718; PubMed-7906398;
                                                                                                                                                                 (TrEMBLrel. 02, (TrEMBLrel. 02, (TrEMBLrel. 08,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ţ
                                                                                                                                     PRELIMINARY;
                                                                                                                                                                (TremBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nature 368:32-38(1994).
                                                                                                                                                                                                                                           Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Waterston R.;
Submitted (AUG-1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (SEP-1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                     NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Simi
Matches 68;
                                                                                                                                                                 01-FEB-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Fulton B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                elegans
                                                                                                                                     Q94185
Q94185;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        191
                                                                                                         RESULT 13
                                                                                                                                                                   쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                           g
                                  ö
```

```
SQGSTSPVYSTISQGSTE -> KEIDQTAINTYKTYFNFAL
LVASKLNNESILTGYIDNFGYSAGLNDHQYYPTDDYNGIKS
VPPRIOGTDDDIDLDKVDNSLATADWTPPVADGTCMFIT
SAAPEDEYGGTTIKGTYTYFFTVVGYLVGGAKSIPGLSIDK
NIVITNNTMNDRDASAVVSKLLELLPTA (IN ISOFORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1572 QMSTSQGS-----SAGSTVASSTAGLVSTSTV-----PSSTGTWGSTSSGTVGS 1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VASSSPAPSTSQNPNPSTSSGSSMI -> LATTSAPKPSVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSQSTSPVESSTSGATSSSGSPGTTLTSISPSPSPSTIGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1402 TSEGSSKASSSPVPSQTSSTPTNPTGSTESSTLLSSTISGSTQH---TTMSKASSGSTSP 1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     337 NAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQTTTT 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------SSTOPOMSTSOGSS---AGSTVASSTTGLVSTSTV------PSS 1684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLEMYD (IN ISOFORM B).
PYPSQSTSPVESSTTPSPGSPGTTLTSTSPSPSQSTTIGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SSGSTVTVGSTEGSSSPIPSTSQNTNPSTSSGSSMSTQTPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BY
IS THAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QGSTSPGISTTSEEMTSQGSTQTPGSTGSTVTQPSTVSDST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56 ADSGDNRASDVP-----GLPVNPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFR 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------MYTVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTET 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              110 VETQEDGKHIAVG--QRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAGHA--- 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
-----TIETTISEPSTIESTIVDISS 844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          277 TATRDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                217 QTSTSTSSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 161;
                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;
Geisel C., Gattung S.;
Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
-!- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) MAY BE PRODUCED ALTERNATIVE PRODUCTS THE SAME GENE. THE SEQUENCE SHOWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 2232;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9 NPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPLGSRALFTPVRNSM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B).
MW; 08D69FA638E14CC8 CRC64;
                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 165.5; DB 5;
Pred. No. 0.12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81; Mismatches 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1616 TISESSTTASA----SSQTGSTVTMGSSSTSGVSTSSA--
                                                                                                                             Ş
                                                                                                                          PRT; 2232
                                                                                                                                                                           Created)
     798 TITITITIEKTISKTITEKPITSESA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 213840
                                                                                                                                                                           03,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U80846; AAC70889.1;
EMBL; U80846; AAC70890.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                             PRELIMINARY;
                                                                                                                                                                        01-MAY-1997 (TrEMBLrel.
01-MAY-1999 (TrEMBLrel.
01-OCT-2000 (TrEMBLrel.
                                                                                                                                                                                                                                                                                             Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alternative splicing.
VARSPLIC 842 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2232 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THE A FORM
                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   870
                                                                                                                                                                                                                                                  K06A9.1 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Sin
Matches 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARSPLIC
```

```
18;
                                            -----QMSTSQGSSAGSTVVSSTASPAASSTAPSST 1768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-MAT-2010 (TrEMBLrel. 16, Last annotation update)
EG:115C2.10 PROTEIN.
EG:115C2.10 OR CG13363.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Metazoa; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSLRSDPKLWLALGTVATGLIGLAATGIVQALALTPEPDSPTTTDPDAAASATETATRDQ 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STASADEVI------APVVAASISLPSKAPVVLMPRCKPAQMAIAALHQ 670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        671 SQÓRQLRRSERQKEKLTD-----GESSD----TSSEQQKK----EQKQQDHQLPQKMFS 716
TTTTTTTSARIVENKPANNIPAQGNVDTPGSEDIMESRRSSMASISSIFFDISSIGGPCR 456
                                                                                                                                                      457 IRMLMLKHRCMIRRCRLLLILIRLFRIWGIQIS------VVYSTIQHPPRDTTDNGA 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 RAS-----DVPGLPVNPMRLA-----ASEITLNDGFEVLHDHGPLDTLNRQ 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |: : | | : | 1| 489 EATAAVSLLEKKLPUVVVSPLTMKELRQKGMTKYDAEMIMANAAYQQQHHH------ 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 IGSSVFRVETQEDGKHIAVGQRNGVETSVVLSDQEYARLQSIDPEGKDKFVFTGGRGGAG 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         540 --QHHFHHHHHHHHHHHHHQHASTGAEATAAVQОМААМQКРG------VGGTGAAG 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      163 HAMVTVASDITEARQRILELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTST 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               589 NAGATIVSSVA-----TSSSI 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           283 LTKEAFQNPDNQKVNIDELGNAIPSGVLKDDVVANIEEQAKAAGEEAKQQ------A 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 NLGHNPNVNNSIPPAPPLPSQTDG---AGGRGQLINSTGPLGSRALFTPVRNSMADSGDN 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Salles C., Valenti P., Darlamitsou A., Henderson N., Campbell L.,
Glover D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Sequencing the distal X chromosome of Drosophila melanogaster."; Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 5.7%; Score 162.5; DB 5; Length 1291; Best Local Similarity 20.6%; Pred. No. 0.088; Matches 99; Conservative 61; Mismatches 189; Indels 131;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Benos P.;
Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
EMBL, AL031581; CAA20894.2; -:
FlyBase; FBgn0025639; EG:115C2.10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Interpro; IPR001214; -.
Pfam; PF00856; SET; 1.
SMART; SM00317; SET; 1.
SEQUENCE 1291 AA; 137366 WW; 021DD566196196BA CRC64;
                                                                                                                                                                                                                                                                                                                                                         1769 GTMSSTTSGTVGSTMSQSSTPARSTTSHTGSTVTLGSSS---TSSN 1810
                                                                                                                                                                                                                                                                                                        507 RLLGNPSAG-IQSTYARLALSGGLRHDMGG-LTGGSNSAVNTSNN 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1291 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
077261
AC 077261;
DT 01-N0V-1
DT 01-N0V-1
DT 01-N0V-1
DE EG:115C;
GN EG:115C;
GN EG:115C;
GN CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHOYOO
CC EPHO
397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              628
                                                                                                                                                                                                                               1738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 223
                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

93	09	20	20
AALHRKNOPVEOT 39	TAHNNNIA 76	ASTSSTFFDTSS 45	SEASSIPSSISS 82
SGALILGGGIGVAVTAALHRKNQI	VARLA	DIPGSEDIMESRRSSN	SNLSSRLSVKSRKPAI
SEELKVSSGAGYGLS	KRVTRNSAGRVGL	SNKPANNTPAQGNVI	WHNSNNSSRINHN
334 IENNAQAQKKYDEQQAKRQEELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVEQT 393	717 LAEEPQPEKS-EEKQQEQQKRVTRNSAGRVGLVARLATAHNNNIA 760	394 TITITITITISARIVENKPANNIPAQGNVDIPGSEDIMESRRSSMASISSIFFDISS 450	761 TTTNSSSSSNKATTITNCNNHNSNNSSRINHNSNLSSRLSVKSRKPAPSEASSIPSSTSS 820
334 I	717 İ	394 T	761 T
δy	qq	οy	QQ

Search completed: September 27, 2001, 14:27:07 Job time: 318 sec

THIS PACE BLANK IUSPIO,